PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00 | A2

(11) International Publication Number: **WO 98/27114**

(43) International Publication Date: 25 June 1998 (25.06.98)

(21) International Application Number: PCT/US97/23321

(22) International Filing Date: 16 December 1997 (16.12.97)

(30) Priority Data:
60/033,601 17 December 1996 (17.12.96) US

(71) Applicant: SCHERING CORPORATION [US/US]; 2000 Galloping Hill Road, Kenilworth, NJ 07033 (US).

(72) Inventor: GORMAN, Daniel, M.; 6371 Central Avenue, Newark, CA 94560 (US).

(74) Agents: THAMPOE, Immac, J. et al.; Schering-Plough Corporation, Patent Dept., K-6-1, 1990, 2000 Galloping Hill Road, Kenilworth, NJ 07033-0530 (US).

(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MAMMALIAN CELL SURFACE ANTIGENS; RELATED REAGENTS

(57) Abstract

Purified genes encoding a T cell surface antigen from a mammal, reagents related thereto including purified proteins, specific antibodies, and nucleic acids encoding this antigen are provided. Methods of using said reagents and diagnostic kits are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

MAMMALIAN CELL SURFACE ANTIGENS; RELATED REAGENTS

FIELD OF THE INVENTION

The present invention pertains to compositions related to proteins which function in controlling activation and expansion of mammalian cells, e.g., cells of a mammalian immune system. In particular, it provides purified genes, proteins, antibodies, and related reagents useful, e.g., to regulate activation, development, differentiation, and function of various cell types, including hematopoietic cells.

BACKGROUND OF THE INVENTION

The activation of resting T cells is critical to most immune responses and allows these cells to exert their regulatory or effector capabilities. See Paul (ed.; 1993) Fundamental Immunology 3d ed., Raven Press, N.Y. Increased adhesion between T cells and antigen presenting cells (APC) or other forms of primary stimuli, e.g., immobilized monoclonal antibodies (mAb), can potentiate the T-cell receptor signals. T-cell activation and T cell expansion depends upon engagement of the T-cell receptor (TCR) and co-stimulatory signals provided by accessory cells. See, e.g., Jenkins and Johnson (1993) Curr. Opin. Immunol. 5:361-367; Bierer and Hahn (1993) Semin. Immunol. 5:249-261; June, et al. (1990) Immunol. Today 11:211-216; and Jenkins (1994) Immunity 1:443-446. A major, and well-studied, co-stimulatory interaction for T cells involves either CD28 or CTLA-4 on T cells with either B7 or B70 (Jenkins (1994) Immunity 1:443-446). Recent studies on CD28 deficient mice (Shahinian, et al. (1993) Science 261:609-612; Green, et al. (1994) Immunity 1:501-508) and CTLA-4 immunoglobulin expressing transgenic mice (Ronchese, et al. (1994) J. Exp. Med. 179:809-817) have revealed deficiencies in some T-cell responses though these mice have normal primary immune responses and normal CTL responses to lymphocytic

choriomeningitis virus and vesicular stomatitis virus. As a result, both these studies conclude that other co-stimulatory molecules must be supporting T-cell function. However, identification of these molecules which mediate distinct costimulatory signals has been difficult.

Tumor Necrosis Factor (TNF) is the prototypic member of an emerging family of cytokines that function as prominent mediators of immune regulation and the inflammatory response. These ligands are typically type II membrane proteins, with homology at the carboxy terminus. A proteolytic processed soluble protein often is produced. See, e.g., Smith, et al. (1994) Cell 76-959-962; Armitage (1994) Current Opinion in Immunology 6:407-413; Gruss and Dower (1995) Blood 85:3378-3404; Wiley, et al. (1995) Immunity 3:673-682; and Baker and Reddy (1996) Oncogene 12:1-9. Crucial roles for these family members are evidenced by a number of studies, and they are implicated in regulation of apoptosis, peripheral tolerance, Ig maturation and isotype switching, and general B cell and T cell functions. See, e.g., Thomson (ed. 1994) The Cytokine Handbook Academic Press, San Diego, CA. These imply fundamental roles in immune and developmental networks.

The inability to modulate activation signals prevents control of inappropriate developmental or physiological responses in the immune system. The present invention provides at least one alternative costimulatory molecule, agonists and antagonists of which will be useful in modulating a plethora of immune responses.

SUMMARY OF THE INVENTION

The present invention is based, in part, upon the discovery of an antigen which exhibits sequence homology to proteins which act as inducers of apoptosis. In particular, it provides a gene encoding a 285 amino acid protein, designated 63954, which is expressed on a number of T cells. Engagement of 63954 may modulate antigen-specific proliferation and cytokine production by effector cells.

63954 is a novel cell surface molecule which, when engaged, may either potentiate immune cell expansion or apoptosis. The human embodiment is described, enabling mammalian genes, proteins, antibodies, and uses thereof. Functional equivalents exhibiting significant sequence homology are available from other mammalian, e.g., human, and non-mammalian species. Moreover, the receptor of 63954 can function as its binding partner to stimulate other cells expressing the receptor.

More particularly, the present invention provides a substantially pure or recombinant 63954 protein or peptide fragment thereof. Various embodiments include a protein or peptide selected from a protein or peptide from a warm blooded animal selected from the group of birds and mammals, including a primate or rodent; a protein or peptide comprising at least one polypeptide segment of SEQ ID NO: 2, 4, 6, or 8; a protein or peptide which exhibits a post-translational modification pattern distinct from natural 63954; or a protein or peptide which is capable of co-stimulating a T cell with another signal. The protein or peptide can comprise a sequence from the extracellular or the intracellular portion of a 63954; or be a fusion protein. The invention further provides a composition of matter selected from: a substantially pure or recombinant 63954 protein or peptide exhibiting at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 2, 4, 6, or 8; a natural sequence 63954 of SEQ ID NO: 2, 4, 6, or 8; or a fusion protein comprising 63954 sequence. In certain preferred embodiments, the substantially pure or isolated protein comprising a segment exhibiting sequence identity to a corresponding portion of a 63954, wherein: said homology is at least about 90% identity and said portion is at least about 9 amino acids; said homology is at least about 80% identity and said portion is at least about 17 amino acids; or said homology is at least about 70% identity and said portion is at least about 25 amino acids. Other embodiments include, e.g., the

composition of matter described, wherein said: 63954 comprises a mature sequence of Table 1; or protein or peptide: is from a warm blooded animal selected from a mammal, including a primate or rodent; comprises at least one polypeptide segment of SEQ ID NO: 2, 4, 6, or 8; exhibits a plurality of portions exhibiting said identity; is a natural allelic variant of 63954; has a length at least about 30 amino acids; exhibits at least two non-overlapping epitopes which are specific for a mammalian 63954; exhibits a sequence identity at least about 90% over a length of at least about 20 amino acids to a primate 63954; exhibits at least two non-overlapping epitopes which are specific for a primate 63954; exhibits a sequence identity at least about 90% over a length of at least about 20 amino acids to a primate 63954; is not glycosylated; is a synthetic polypeptide; is attached to a solid substrate; is conjugated to another chemical moiety; is a 5-fold or less substitution from natural sequence; or is a deletion or insertion variant from a natural sequence. Other embodiments include a composition comprising: a sterile 63954 protein or peptide; or said 63954 protein or peptide and a carrier, wherein said carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for oral, rectal, nasal, topical, or parenteral administration. Fusion protein forms include those comprising: mature protein comprising sequence of Table 1; a detection or purification tag, including a FLAG, His6, or Ig sequence; or sequence of another TNF ligand protein. Kits include, e.g., those comprising said protein or polypeptide, and: a compartment comprising said protein or polypeptide; and/or instructions for use or disposal of reagents in said kit.

Another embodiment is a composition comprising a 63954 protein and a pharmaceutically acceptable carrier. Other compositions may combine said entities with an agonist or antagonist of other T cell signaling molecules, e.g., signaling entities through the T cell receptor, CD40, CD40 ligand, CTLA-8, CD28, B7, B70, BAS-1, SLAM, etc.

The invention also embraces an antibody which specifically binds a 63954 protein or peptide, e.g., wherein the 63954 is a mammalian protein, including a human; the antibody is raised against a purified 63954 peptide sequence of SEQ ID NO: 2, 4, 6, or 8; the antibody is a monoclonal antibody; or the antibody is labeled. Other binding compounds are provided, e.g., comprising an antigen binding portion from an antibody, which specifically binds to a natural 63954 protein, wherein: said protein is a primate protein; said binding compound is an Fv, Fab, or Fab2 fragment; said binding compound is conjugated to another chemical moiety; or said antibody: is raised against a peptide sequence of a mature polypeptide comprising sequence of Table 1; is raised against a mature 63954; is raised to a purified 63954; is immunoselected; is a polyclonal antibody; binds to a denatured 63954; exhibits a Kd to antigen of at least 30 μM; is attached to a solid substrate, including a bead or plastic membrane; is in a sterile composition; or is detectably labeled, including a radioactive or fluorescent label. Kits include, e.g., those comprising said binding compound, and: a compartment comprising said binding compound; and/or instructions for use or disposal of reagents in said kit.

Such binding compositions also provide methods of purifying a 63954 protein or peptide from other materials in a mixture comprising contacting said mixture to an antibody, and separating bound 63954 from other materials;

Certain other compositions include those comprising: a sterile binding compound, or said binding compound and a carrier, wherein said carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for oral, rectal, nasal, topical, or parenteral administration.

Another aspect of the invention is an isolated or recombinant nucleic acid capable of encoding a 63954 protein or peptide, including a nucleic acid which encodes a sequence of SEQ ID NO: 2, 4, 6, or 8; which includes a

sequence of SEQ ID NO: 1, 3, 5, or 7; which encodes a sequence from an extracellular domain of a natural 63954; or which encodes a sequence from an intracellular domain of a natural 63954. Such nucleic acid embodiments also include an expression or replicating vector. Various other nucleic acid embodiments are provided, e.g., an isolated or recombinant nucleic acid encoding said protein or peptide or fusion protein, wherein: said TNF ligand family protein is from a mammal, including a primate; or said nucleic acid: encodes an antigenic peptide sequence of Table 1; encodes a plurality of antigenic peptide sequences of Table 1; exhibits at least about 80% identity to a natural cDNA encoding said segment; is an expression vector; further comprises an origin of replication; is from a natural source; comprises a detectable label; comprises synthetic nucleotide sequence; is less than 6 kb, preferably less than 3 kb; is from a mammal, including a primate; comprises a natural full length coding sequence; is a hybridization probe for a gene encoding said TNF ligand family protein; or is a PCR primer, PCR product, or mutagenesis primer. The invention also provides a cell or tissue comprising such a recombinant nucleic acid, e.g., wherein said cell is: a prokaryotic cell; a eukaryotic cell; a bacterial cell; a yeast cell; an insect cell; a mammalian cell; a mouse cell; a primate cell; or a human cell.

Also provided are a method of expressing a 63954 peptide by expressing a nucleic acid encoding a 63954 polypeptide. The invention also provides a cell, tissue, organ, or organism comprising a nucleic acid encoding a 63954 peptide.

Kit embodiments include those, e.g., which comprise said nucleic acid and: a compartment further comprising a 63954 protein or polypeptide; and/or instructions for use or disposal of reagents in said kit.

The invention further provides a nucleic acid which: hybridizes under wash conditions of 30° C and less than 2M salt to SEQ ID NO: 1, 3, 5, or 7; or exhibits at least about

85% identity over a stretch of at least about 30 nucleotides to a mammalian 63954, including a primate or rodent. In other embodiments, the nucleic acid hybridizes where the nucleic acid, wherein: said wash conditions are at 45° C and/or 500 mM salt; or said identity is at least 90% and/or said stretch is at least 55 nucleotides. In yet other embodiments, the nucleic acid hybridizes, wherein: said wash conditions are at 55° C and/or 150 mM salt; or said identity is at least 95% and/or said stretch is at least 75 nucleotides.

The invention also provides a kit containing a substantially pure 63954 or fragment; an antibody or receptor which specifically binds a 63954; or a nucleic acid, or its complement, encoding a 63954 or peptide. This kit also provides methods for detecting in a sample the presence of a nucleic acid, protein, or antibody, comprising testing said sample with such a kit.

The invention also provides a recombinant nucleic acid comprising sequence at least about 70% identity over a stretch of at least about 30 nucleotides to a 63954 nucleic acid sequence of SEQ ID NO: 1, 3, 5, or 7, useful, e.g., as a probe or PCR primer for a related gene. Another embodiment further encodes a polypeptide comprising at least about 60% identity over a stretch of at least about 20 amino acids to a 63954 sequence of SEQ ID NO: 2, 4, 6, or 8.

The invention also supplies methods of modulating the physiology of a cell comprising contacting said cell with a substantially pure 63954 or fragment; an antibody or binding partner which specifically binds a 63954; or a nucleic acid encoding a 63954 or peptide. Certain preferred embodiments include a method where the cell is a T cell and the modulating of physiology is activation of the T cell or apoptosis of the T cell; or where the cell is in a tissue and/or in an organism.

Another method provided is treating an organism having an abnormal immune response by administering to said organism an effective dose of: an antibody or binding

partner which binds specifically to a 63954; a substantially pure 63954 protein, or peptide thereof; or a nucleic acid encoding a 63954 peptide. The abnormal immune response may be characterized by a T cell immune deficiency; chronic inflammation; or tissue rejection.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

All references cited herein are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

I. General

The present invention provides amino acid sequences and DNA sequences encoding various mammalian proteins which are antigens found on many T cells. Among these proteins are antigens which modulate, e.g., induce or prevent proliferation or differentiation of interacting cells, among other physiological effects. The full length antigens, and fragments, or antagonists will be useful in physiological modulation of cells expressing counter receptors for the antigen. The proteins will also be useful as antigens, e.g., immunogens, for raising antibodies to various epitopes on the protein, both linear and conformational epitopes. The molecule may be useful in defining functional T cell or NK cell subsets.

63954 exhibits structural motifs characteristic of a member of the TNF ligand family. Compare, e.g., with the CD40 ligand, OX40 ligand, TNF, NGF, and FAS. Table 1 illustrates the nucleic acid (SEQ ID NO: 1, 3, 5, and 7) and predicted amino acid sequences (SEQ ID NO: 2, 4, 6, and 8) for human 63954 variants, and for a closely related mouse gene.

Table 1: Human 63954 nucleotide sequence (SEQ ID NO: 1), with an open reading frame running from about base pairs 157 through 1011. Nucleotide 10 is designated C, but may be A, C, G, or T. Predicted intracellular domain sequence runs about from met1 to gly49; residues 124 and 242 are potential glycosylation sites; a transmembrane sequence probably runs about from cys50 to pro68; and the extracellular domain probably runs about from gly69 to leu285.

```
GGGCAGAGAC AAATTCAGGA TAACTCTCCT GAGGGGTGAG CCAAGCCCTG CCATGTAGTG      60

CACGCAGGAC ATCAACAAAC ACAGATAACA GGAAATAATC CATTCCCTGT GGTCACTTAT     120

TCTAAAGGCC CCAACCTTCA AAGTTCAAGT AGTGAT ATG GAT GAC TCC ACA GAA       174
                                        Met Asp Asp Ser Thr Glu
                                         1               5

AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA GAA GAA ATG       222
Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg Glu Glu Met
            10                  15                  20

AAA CTG CAA GGA GTG TGT TTC CAT CCT CCC ACG GAA GGA AAG CCC CTC       270
Lys Leu Gln Gly Val Cys Phe His Pro Pro Thr Glu Gly Lys Pro Leu
        25                  30                  35

TGT CCG ATC CTC CAA AGA CGG AAA GCT GCT GGC TGC AAC CTT GCT GCT       318
Cys Pro Ile Leu Gln Arg Arg Lys Ala Ala Gly Cys Asn Leu Ala Ala
    40                  45                  50

GGC ACT GCT GTC TTG CTG CCT CAC GGT GGT GTC TTT CTA CCA GGT GGC       366
Gly Thr Ala Val Leu Leu Pro His Gly Gly Val Phe Leu Pro Gly Gly
55                  60                  65                  70

CGC CCT GCA AGG GAC CTG GCC AGC CTC CGG GCA GAG CTG CAG GGC CAC       414
Arg Pro Ala Arg Asp Leu Ala Ser Leu Arg Ala Glu Leu Gln Gly His
                75                  80                  85

CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG GCC GGC CTG       462
His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys Ala Gly Leu
            90                  95                  100

GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT GAA CCA CCA       510
Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe Glu Pro Pro
        105                 110                 115

GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT AAG CGT GCC       558
Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn Lys Arg Ala
    120                 125                 130

GTT CAG GGT CCA GAA GAA ACA GTC ACT CAA GAC TGC TTG CAA CTG ATT       606
Val Gln Gly Pro Glu Glu Thr Val Thr Gln Asp Cys Leu Gln Leu Ile
135                 140                 145                 150
```

```
GCA GAC AGT GAA ACA CCA ACT ATA CAA AAA GGA TCT TAC ACA TTT GTT      654
Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys Gly Ser Tyr Thr Phe Val
                155                 160                 165

CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA GAA AAA GAG      702
Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu Glu Lys Glu
            170                 175                 180

AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA TAT GGT CAG      750
Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile Tyr Gly Gln
        185                 190                 195

GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA GTT CAG AGG      798
Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu Val Gln Arg
    200                 205                 210

AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG ACT TTG TTT      846
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
215                 220                 225                 230

CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT TCC TGC TAT      894
Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn Ser Cys Tyr
                235                 240                 245

TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC CAA CTT GCA      942
Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu Gln Leu Ala
            250                 255                 260

ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT GTC ACA TTT      990
Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp Val Thr Phe
        265                 270                 275

TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC TGTAGCTATT        1041
Phe Gly Ala Leu Lys Leu Leu
    280                 285

TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC CAAAAAAAAA  1101

AAAAAAAA                                                           1109
```

A frameshift resulting from the deletion of the nucleotide C at position 229, and insertion of a nucleotide G between positions 377 and 378 results in a changed transcript and protein, as found (SEQ ID NO: 3 and 4):

```
CACAAAGGCA CAAAGGAGAA AATTCAGGAT AACTCTCCTG AGGGGTGAGC CAAGCCCTGC   60

CATGTATTGC ACGCAGGACA TCAACAAACA CAGATAACAG GAAATGATCC ATTCCCTGTG  120

GTCACTTATT CTAAAGGCCC CAACCTTCAA AGTTCAAGTA GTGAT ATG GAT GAC      174
                                                 Met Asp Asp
                                                   1
```

```
TCC ACA GAA AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA        222
Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg
  5                  10                  15

GAA GAA ATG AAA CTG AAG GAG TGT GTT TCC ATC CTC CCA CGG AAG GAA        270
Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro Arg Lys Glu
 20                  25                  30                  35

AGC CCC TCT GTC CGA TCC TCC AAA GAC GGA AAG CTG CTG GCT GCA ACC        318
Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu Ala Ala Thr
                 40                  45                  50

TTG CTG CTG GCA CTG CTG TCT TGC TGC CTC ACG GTG GTG TCT TTC TAC        366
Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val Ser Phe Tyr
             55                  60                  65

CAG GTG GCC GCC CTG CAA GGG GAC CTG GCC AGC CTC CGG GCA GAG CTG        414
Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg Ala Glu Leu
         70                  75                  80

CAG GGC CAC CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG        462
Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys
  85                  90                  95

GCC GGC CTG GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT        510
Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe
100                 105                 110                 115

GAA CCA CCA GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT        558
Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn
                120                 125                 130

AAG CGT GCC GTT CAG GGT CCA GAA GAA ACA GTC ACT CAA GAC TGC TTG        606
Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Val Thr Gln Asp Cys Leu
            135                 140                 145

CAA CTG ATT GCA GAC AGT GAA ACA CCA ACT ATA CAA AAA GGA TCT TAC        654
Gln Leu Ile Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys Gly Ser Tyr
        150                 155                 160

ACA TTT GTT CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA        702
Thr Phe Val Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu
    165                 170                 175

GAA AAA GAG AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA        750
Glu Lys Glu Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile
180                 185                 190                 195

TAT GGT CAG GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA        798
Tyr Gly Gln Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu
                200                 205                 210
```

```
ATT CAG AGG AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG        846
Ile Gln Arg Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val
            215                 220                 225

ACT TTG TTT CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT        894
Thr Leu Phe Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn
        230                 235                 240

TCC TGC TAT TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC        942
Ser Cys Tyr Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu
    245                 250                 255

CAA CTT GCA ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT        990
Gln Leu Ala Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp
260                 265                 270                 275

GTC ACA TTT TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC         1040
Val Thr Phe Phe Gly Ala Leu Lys Leu Leu
                280                 285

TGTAGCTATT TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC     1100

C                                                                     1101
```

An alternative form (perhaps due to alternate splicing) resulting from the deletion of nucleotides 589-645 of SEQ ID NO: 3 results in another form (SEQ ID NO: 5 and 6):

```
CCCCTTTGGG TCGCCCACTA ATCCGAACAA ACCCCCATAA AACGAAATGA TCCATTCCCT      60

GTGGTCACTT ATTCTAAAGG CCCCAACCTT CAAAGTTCAA GTAGTGAT ATG GAT GAC      117
                                                    Met Asp Asp
                                                     1

TCC ACA GAA AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA        165
Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg
      5                  10                  15

GAA GAA ATG AAA CTG AAG GAG TGT GTT TCC ATC CTC CCA CGG AAG GAA        213
Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro Arg Lys Glu
 20                  25                  30                  35

AGC CCC TCT GTC CGA TCC TCC AAA GAC GGA AAG CTG CTG GCT GCA ACC        261
Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu Ala Ala Thr
                40                  45                  50

TTG CTG CTG GCA CTG CTG TCT TGC TGC CTC ACG GTG GTG TCT TTC TAC        309
Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val Ser Phe Tyr
            55                  60                  65

CAG GTG GCC GCC CTG CAA GGG GAC CTG GCC AGC CTC CGG GCA GAG CTG        357
Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg Ala Glu Leu
        70                  75                  80
```

```
CAG GGC CAC CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG        405
Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys
     85                  90                  95

GCC GGC CTG GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT        453
Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe
100                 105                 110                 115

GAA CCA CCA GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT        501
Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn
                120                 125                 130

AAG CGT GCC GTT CAG GGT CCA GAA GAA ACA GGA TCT TAC ACA TTT GTT        549
Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Gly Ser Tyr Thr Phe Val
            135                 140                 145

CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA GAA AAA GAG        597
Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu Glu Lys Glu
        150                 155                 160

AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA TAT GGT CAG        645
Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile Tyr Gly Gln
    165                 170                 175

GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA ATT CAG AGG        693
Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu Ile Gln Arg
180                 185                 190                 195

AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG ACT TTG TTT        741
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
                200                 205                 210

CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT TCC TGC TAT        789
Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn Ser Cys Tyr
            215                 220                 225

TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC CAA CTT GCA        837
Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu Gln Leu Ala
        230                 235                 240

ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT GTC ACA TTT        885
Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp Val Thr Phe
    245                 250                 255

TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC TGTAGCTATT           936
Phe Gly Ala Leu Lys Leu Leu
260                 265

TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC C               987
```

Mouse EST GenBank Accession number AA254047 (SEQ ID NO: 7 asnd 8):

```
TC TCA GCT CCT CCT GCA CCA TGC CTG CCT GGA TGC CGC CAT TCT CAA          47
   Ser Ala Pro Pro Ala Pro Cys Leu Pro Gly Cys Arg His Ser Gln
     1               5                  10                  15

CAT GAT GAT AAT GGA ATG AAC CTC AGA AAC AGA ACT TAC ACA TTT GTT          95
His Asp Asp Asn Gly Met Asn Leu Arg Asn Arg Thr Tyr Thr Phe Val
                 20                  25                  30

CCA TGG CTT CTC AGC TTT AAA AGA GGA AAT GCC TTG GAG GAG AAA GAG         143
Pro Trp Leu Leu Ser Phe Lys Arg Gly Asn Ala Leu Glu Glu Lys Glu
             35                  40                  45

AAC AAA ATA GTG GTG AGG CAA ACA GGC TAT TTC TTC ATC TAC AGC CAG         191
Asn Lys Ile Val Val Arg Gln Thr Gly Tyr Phe Phe Ile Tyr Ser Gln
         50                  55                  60

GTT CTA TAC ACG GAC CCC ATC TTT GCT ATG GGT CAT GTC ATC CAG AGG         239
Val Leu Tyr Thr Asp Pro Ile Phe Ala Met Gly His Val Ile Gln Arg
     65                  70                  75

AAG AAA GTA CAC GTC TTT GGG GAC GAG CTG AGC CTG GTG ACC CTG TTC         287
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
 80                  85                  90                  95

CGA TGT ATT CAG AAT ATG CCC AAA ACA CTG CCC AAC AAT TCC TGC TAC         335
Arg Cys Ile Gln Asn Met Pro Lys Thr Leu Pro Asn Asn Ser Cys Tyr
                100                 105                 110

TCG GCT GGC ATC GCG AGG CTG GAA GAA GGA GAT GAG ATT CAG CTT GCA         383
Ser Ala Gly Ile Ala Arg Leu Glu Glu Gly Asp Glu Ile Gln Leu Ala
            115                 120                 125

ATT CCT CGG GAG AAT GCA CAG ATT TCA CGC AAC GGA GAC GAC ACC TTC         431
Ile Pro Arg Glu Asn Ala Gln Ile Ser Arg Asn Gly Asp Asp Thr Phe
        130                 135                 140

TTT GGT GCC CTA AAA CTG CTG TAACTCACTT GCTGGAGTGC GTGATCCCCT            482
Phe Gly Ala Leu Lys Leu Leu
    145                 150

TCCCTCGTCT TCTCTGTACC TCCGAGGGAG AAACAGACGA CTGGAAAAAC TAAAAGATGG       542

GGAAAGCCGT CAGCGAAAGT TTTCTCGTGA CCCGTTGAAT CTGATCCAAA CCAGGAAATA       602

TAACAGACAG CCACA                                                       617
```

Table 1 (continued): Alignment of the different forms:

```
SEQ ID NO:2:  MDDSTEREQS RLTSCLKKRE EMKLQGVCFH PPTEGKPLCP ILQRRKAAGC
SEQ ID NO:4:  MDDSTEREQS RLTSCLKKRE EMKLKECVSI LPRKESPSVR SSKDGKLLAA
SEQ ID NO:6:  MDDSTEREQS RLTSCLKKRE EMKLKECVSI LPRKESPSVR SSKDGKLLAA

SEQ ID NO: 2:  NLAAGTAVLL PHGGVFLPGG RPARDLASLR AELQGHHAEK LPAGAGAPKA
SEQ ID NO: 4:  TLLLALLSCC LTVVSFYQVA ALQGDLASLR AELQGHHAEK LPAGAGAPKA
SEQ ID NO: 6:  TLLLALLSCC LTVVSFYQVA ALQGDLASLR AELQGHHAEK LPAGAGAPKA

SEQ ID NO: 2:  GLEEAPAVTA GLKIFEPPAP GEGNSSQNSR NKRAVQGPEE TVTQDCLQLI
SEQ ID NO: 4:  GLEEAPAVTA GLKIFEPPAP GEGNSSQNSR NKRAVQGPEE TVTQDCLQLI
SEQ ID NO: 6:  GLEEAPAVTA GLKIFEPPAP GEGNSSQNSR NKRAVQGPEE T---------
SEQ ID NO: 8:              ----SAPPAP CLPGCRHSQH DDNGMNLRN  ----------

SEQ ID NO: 2:  ADSETPTIQK GSYTFVPWLL SFKRGSALEE KENKILVKET GYFFIYGQVL
SEQ ID NO: 4:  ADSETPTIQK GSYTFVPWLL SFKRGSALEE KENKILVKET GYFFIYGQVL
SEQ ID NO: 6:  ---------- GSYTFVPWLL SFKRGSALEE KENKILVKET GYFFIYGQVL
SEQ ID NO: 8:  ---------- RTYTFVPWLL SFKRGNALEE KENKIVVRQT GYFFIYSQVL

SEQ ID NO: 2:  YTDKTYAMGH LVQRKKVHVF GDELSLVTLF RCIQNMPETL PNNSCYSAGI
SEQ ID NO: 4:  YTDKTYAMGH LIQRKKVHVF GDELSLVTLF RCIQNMPETL PNNSCYSAGI
SEQ ID NO: 6:  YTDKTYAMGH LIQRKKVHVF GDELSLVTLF RCIQNMPETL PNNSCYSAGI
SEQ ID NO: 8:  YTDPIFAMGH VIQRKKVHVF GDELSLVTLF RCIQNMPKTL PNNSCYSAGI

SEQ ID NO: 2:  AKLEEGDELQ LAIPRENAQI SLDGDVTFFG ALKLL
SEQ ID NO: 4:  AKLEEGDELQ LAIPRENAQI SLDGDVTFFG ALKLL
SEQ ID NO: 6:  AKLEEGDELQ LAIPRENAQI SLDGDVTFFG ALKLL
SEQ ID NO: 8:  ARLEEGDEIQ LAIPRENAQI SRNGDDTFFG ALKLL
```

TNF ligand family members have a conserved leucine residue corresponding to 186 (residues corresponding to SEQ ID NO: 2); a conserved glycine residue corresponding to residue 191; a conserved tyrosine residue corresponding to 196; a conserved glycine residue corresponding to residue 249; a conserved leucine residue corresponding to residue 253; a conserved phenylalanine residue corresponding to 279; and a conserved glycine residue corresponding to residue 280. The TNF ligand domain seems to run about from 186 (leu) to 285 (leu). Related family members include ligands for CD40 and FAS, and lymphotoxin beta, tumor necrosis factor, etc.

This clone was assembled through the careful analysis of ESTs present in various databases, e.g., Merck-WashU public database. The ESTs were identified from several different libraries derived from, e.g., human B cell lymphoma, human bone marrow, PHA activated T-cells, CD34 depleted cord blood, primary dendritic cells, human T-cell lymphoma, macrophage-oxLDL, bone marrow cell line RS4;11, human neutrophil, smooth muscle, stomach cancer, and Soares fetal liver. 63954 should be expressed in these and similar cells and tissues. Other isolates have been characterized, including a mouse counterpart.

The structural homology of 63954 to the TNF ligand family suggests function of this molecule. 63954, as a T cell surface molecule, likely modulates Ag-specific proliferative responses on effector cells, or induction of apoptosis of those cells. 63954 agonists, or antagonists, may also act as a co-stimulatory molecule for regulation of T cell mediated cell activation, and may in fact, cause a shift of T helper cell types, e.g., between Th1 and Th2. Thus, 63954 or antagonists should be useful in the treatment of abnormal immune disorders, e.g., T cell immune deficiencies, chronic inflammation, or tissue rejection.

TNF ligand molecules typically modulate cell proliferation, viability, and differentiation. For example, TNF and FAS can kill cells expressing their respective receptors, including fibroblasts, liver cells, and lymphocytes. Some members of this class of ligands exhibit effects on cellular proliferation of cells expressing their respective receptors, e.g., B cells expressing CD40. These effects on proliferation may also effect subsequent differentiation steps, and may lead, directly or indirectly, to changes in cytokine expression profiles.

The members of the TNF ligand family also exhibit costimulation effects, which may also regulate cellular differentiation or apoptosis. Receptor expressing cells may be protected from activation induced cell death (AICD) or

apoptosis. For example, CD40 ligand can have effects on T and B lymphocytes.

The embodiment characterized herein is from human, but additional sequences for proteins in other mammalian species, e.g., primates and rodents, will also be available. See below. The descriptions below are directed, for exemplary purposes, to a human 63954, but are likewise applicable to related embodiments from other species.

The human 63954 protein is a protein which exhibits structural features characteristic of a cell surface antigen, e.g., a TNF ligand family member. The protein is easily detected on particular cell types, others express lesser amounts. The 63954 antigen should be present in the identified tissue types and the interaction of the antigen with its binding partner should be important for mediating various aspects of cellular physiology or development, as described.

II. Purified 63954

Human 63954 amino acid sequence is shown in SEQ ID NO: 2, 4, and 6. These amino acid sequences, provided amino to carboxy, are important in providing sequence information in the antigen allowing for distinguishing the protein from other proteins and exemplifying numerous variants. Moreover, the peptide sequences allow preparation of peptides to generate antibodies to recognize such segments, and nucleotide sequences allow preparation of oligonucleotide probes, both of which are strategies for detection or isolation, e.g., cloning, of genes or cDNAs encoding such sequences.

As used herein, the term "human 63954" shall encompass, when used in a protein context, a protein having amino acid sequence shown in SEQ ID NO: 4 or 6, or a significant fragment of such a protein, or another highly homologous protein derived from human. These binding components, e.g., antibodies, typically bind to a 63954 with high affinity, e.g., at least about 100 nM, usually better

than about 30 nM, preferably better than about 10 nM, and more preferably at better than about 3 nM. Homologous proteins would be found in mammalian species other than human, e.g., primates or rodents. Non-mammalian species should also possess structurally or functionally related genes and proteins, e.g., birds or amphibians.

The term "polypeptide" as used herein includes a significant fragment or segment, and encompasses a stretch of amino acid residues of at least about 8 amino acids, generally at least about 12 amino acids, typically at least about 16 amino acids, preferably at least about 20 amino acids, and, in particularly preferred embodiments, at least about 30 or more amino acids.

The term "binding composition" refers to molecules that bind with specificity to 63954, e.g., in a cell adhesion pairing type fashion, or an antibody-antigen interaction. It also includes compounds, e.g., proteins, which specifically associate with 63954, including in a natural physiologically relevant protein-protein interaction, either covalent or non-covalent. The molecule may be a polymer, or chemical reagent. A functional analog may be an antigen with structural modifications, or it may be a molecule which has a molecular shape which interacts with the appropriate binding determinants. The compounds may serve as agonists or antagonists of the binding interaction, see, e.g., Goodman, et al. (eds.) (1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics (8th ed.), Pergamon Press.

Substantially pure typically means that the protein is free from other contaminating proteins, nucleic acids, or other biologicals derived from the original source organism. Purity may be assayed by standard methods, typically by weight, and will ordinarily be at least about 40% pure, generally at least about 50% pure, often at least about 60% pure, typically at least about 80% pure, preferably at least about 90% pure, and in most preferred embodiments, at least about 95% pure. Carriers or excipients will often be added.

Solubility of a polypeptide or fragment depends upon the environment and the polypeptide. Many parameters affect polypeptide solubility, including temperature, electrolyte environment, size and molecular characteristics of the polypeptide, and nature of the solvent. Typically, the temperature at which the polypeptide is used ranges from about 4° C to about 65° C. Usually the temperature at use is greater than about 18° C. For diagnostic purposes, the temperature will usually be about room temperature or warmer, but less than the denaturation temperature of components in the assay. For therapeutic purposes, the temperature will usually be body temperature, typically about 37° C for humans and mice, though under certain situations the temperature may be raised or lowered in situ or in vitro.

The size and structure of the polypeptide should generally be in a substantially stable state, and usually not in a denatured state. The polypeptide may be associated with other polypeptides in a quaternary structure, e.g., to confer solubility, or associated with lipids or detergents in a manner which approximates natural lipid bilayer interactions.

The solvent and electrolytes will usually be a biologically compatible buffer, of a type used for preservation of biological activities, and will usually approximate a physiological aqueous solvent. Usually the solvent will have a neutral pH, typically between about 5 and 10, and preferably about 7.5. On some occasions, one or more detergents will be added, typically a mild non-denaturing one, e.g., CHS (cholesteryl hemisuccinate) or CHAPS (3-[3-cholamidopropyl)dimethylammonio]-1-propane sulfonate), or a low enough concentration as to avoid significant disruption of structural or physiological properties of the protein.

III. Physical Variants

This invention also encompasses proteins or peptides having substantial amino acid sequence identity with the amino acid sequence of the 63954. The variants include species, polymorphic, or allelic variants.

Amino acid sequence homology, or sequence identity, is determined by optimizing residue matches, if necessary, by introducing gaps as required. See also Needleham, et al. (1970) J. Mol. Biol. 48:443-453; Sankoff, et al. (1983) Chapter One in Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, Addison-Wesley, Reading, MA; and software packages from IntelliGenetics, Mountain View, CA; and the University of Wisconsin Genetics Computer Group, Madison, WI. Sequence identity changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Homologous amino acid sequences are typically intended to include natural polymorphic or allelic and interspecies variations in each respective protein sequence. Typical homologous proteins or peptides will have from 25-100% identity (if gaps can be introduced), to 50-100% identity (if conservative substitutions are included) with the amino acid sequence of the 63954. Identity measures will be at least about 35%, generally at least about 40%, often at least about 50%, typically at least about 60%, usually at least about 70%, preferably at least about 80%, and more preferably at least about 90%.

The isolated 63954 DNA can be readily modified by nucleotide substitutions, nucleotide deletions, nucleotide insertions, and inversions of nucleotide stretches. These modifications result in novel DNA sequences which encode these antigens, their derivatives, or proteins having similar physiological, immunogenic, antigenic, or other

functional activity. These modified sequences can be used to produce mutant antigens or to enhance expression. Enhanced expression may involve gene amplification, increased transcription, increased translation, and other mechanisms. "Mutant 63954" encompasses a polypeptide otherwise falling within the sequence identity definition of the 63954 as set forth above, but having an amino acid sequence which differs from that of 63954 as normally found in nature, whether by way of deletion, substitution, or insertion. This generally includes proteins having significant identity with a protein having sequence of SEQ ID NO: 4 or 6, and as sharing various biological activities, e.g., antigenic or immunogenic, with those sequences, and in preferred embodiments contain most of the full length disclosed sequences. Full length sequences will typically be preferred, though truncated versions, e.g., soluble constructs and intact domains, will also be useful, likewise, genes or proteins found from natural sources are typically most desired. Similar concepts apply to different 63954 proteins, particularly those found in various warm blooded animals, e.g., mammals and birds. These descriptions are generally meant to encompass all 63954 proteins, not limited to the particular human embodiment specifically discussed.

63954 mutagenesis can also be conducted by making amino acid insertions or deletions. Substitutions, deletions, insertions, or any combinations may be generated to arrive at a final construct. Insertions include amino- or carboxy- terminal fusions. Random mutagenesis can be conducted at a target codon and the expressed mutants can then be screened for the desired activity. Methods for making substitution mutations at predetermined sites in DNA having a known sequence are well known in the art, e.g., by M13 primer mutagenesis or polymerase chain reaction (PCR) techniques. See, e.g., Sambrook, et al. (1989); Ausubel, et al. (1987 and Supplements); and Kunkel, et al. (1987) Methods in Enzymol. 154:367-382.

The present invention also provides recombinant proteins, e.g., heterologous fusion proteins using segments from these proteins. A heterologous fusion protein is a fusion of proteins or segments which are naturally not normally fused in the same manner. A similar concept applies to heterologous nucleic acid sequences. Fusion proteins will be useful as sources for cleaving, separating, and purifying portions thereof.

In addition, new constructs may be made from combining similar functional domains from other proteins. For example, target-binding or other segments may be "swapped" between different new fusion polypeptides or fragments. See, e.g., Cunningham, et al. (1989) Science 243:1330-1336; and O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992.

The phosphoramidite method described by Beaucage and Carruthers (1981) Tetra. Letts. 22:1859-1862, will produce suitable synthetic DNA fragments. A double stranded fragment will often be obtained either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence, e.g., PCR techniques.

IV. Functional Variants

The blocking of physiological response to 63954s may result from the inhibition of binding of the antigen to its binding partner, e.g., another of itself, likely through competitive inhibition. Thus, in vitro assays of the present invention will often use isolated protein, membranes from cells expressing a membrane associated recombinant 63954, soluble fragments comprising antigen binding segments of these proteins, or fragments attached to solid phase substrates. These assays will also allow for the diagnostic determination of the effects of either binding segment mutations and modifications, or antigen mutations and modifications, e.g., 63954 analogs.

This invention also contemplates the use of competitive drug screening assays, e.g., where neutralizing antibodies to antigen or binding fragments compete with a test compound for binding to the protein, e.g., of natural protein sequence.

"Derivatives" of 63954 antigens include amino acid sequence mutants from naturally occurring forms, glycosylation variants, and covalent or aggregate conjugates with other chemical moieties. Covalent derivatives can be prepared by linkage of functionalities to groups which are found in 63954 amino acid side chains or at the N- or C-termini, e.g., by standard means. See, e.g., Lundblad and Noyes (1988) Chemical Reagents for Protein Modification, vols. 1-2, CRC Press, Inc., Boca Raton, FL; Hugli (ed.) (1989) Techniques in Protein Chemistry, Academic Press, San Diego, CA; and Wong (1991) Chemistry of Protein Conjugation and Cross Linking, CRC Press, Boca Raton, FL.

In particular, glycosylation alterations are included, e.g., made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing, or in further processing steps. See, e.g., Elbein (1987) Ann. Rev. Biochem. 56:497-534. Also embraced are versions of the peptides with the same primary amino acid sequence which have other minor modifications, including phosphorylated amino acid residues, e.g., phosphotyrosine, phosphoserine, or phosphothreonine.

Fusion polypeptides between 63954s and other homologous or heterologous proteins are also provided. Many cytokine receptors or other surface proteins are multimeric, e.g., homodimeric entities, and a repeat construct may have various advantages, including lessened susceptibility to proteolytic cleavage. Typical examples are fusions of a reporter polypeptide, e.g., luciferase, with a segment or domain of a protein, e.g., a receptor-binding segment, so that the presence or location of the fused ligand may be easily determined. See, e.g., Dull, et al., U.S. Patent No. 4,859,609. Other gene fusion partners include bacterial ß-

galactosidase, trpE, Protein A, ß-lactamase, alpha amylase, alcohol dehydrogenase, yeast alpha mating factor, and detection or purification tags such as a FLAG sequence of His6 sequence. See, e.g., Godowski, et al. (1988) Science 241:812-816.

Fusion peptides will typically be made by either recombinant nucleic acid methods or by synthetic polypeptide methods. Techniques for nucleic acid manipulation and expression are described generally, e.g., in Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.), vols. 1-3, Cold Spring Harbor Laboratory; and Ausubel, et al. (eds.) (1993) Current Protocols in Molecular Biology, Greene and Wiley, NY. Techniques for synthesis of polypeptides are described, e.g., in Merrifield (1963) J. Amer. Chem. Soc. 85:2149-2156; Merrifield (1986) Science 232: 341-347; Atherton, et al. (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford; and Grant (1992) Synthetic Peptides: A User's Guide, W.H. Freeman, NY.

This invention also contemplates the use of derivatives of 63954s other than variations in amino acid sequence or glycosylation. Such derivatives may involve covalent or aggregative association with chemical moieties. Covalent or aggregative derivatives will be useful as immunogens, as reagents in immunoassays, or in purification methods such as for affinity purification of binding partners, e.g., other antigens. A 63954 can be immobilized by covalent bonding to a solid support such as cyanogen bromide-activated SEPHAROSE, by methods which are well known in the art, or adsorbed onto polyolefin surfaces, with or without glutaraldehyde cross-linking, for use in the assay or purification of anti-63954 antibodies or an alternative binding composition. The 63954s can also be labeled with a detectable group, e.g., for use in diagnostic assays. Purification of 63954 may be effected by an immobilized antibody or complementary binding partner.

A solubilized 63954 or fragment of this invention can be used as an immunogen for the production of antisera or antibodies specific for binding to the antigen or fragments thereof. Purified antigen can be used to screen monoclonal antibodies or antigen-binding fragments, encompassing antigen binding fragments of natural antibodies, e.g., Fab, Fab', F(ab)2, etc. Purified 63954s can also be used as a reagent to detect antibodies generated in response to the presence of elevated levels of the antigen or cell fragments containing the antigen, both of which may be diagnostic of an abnormal or specific physiological or disease condition. This invention contemplates antibodies raised against amino acid sequences encoded by nucleotide sequence shown in SEQ ID NO: 1, 3, 5, or 7, or fragments of proteins containing it. In particular, this invention contemplates antibodies having binding affinity to or being raised against specific fragments which are predicted to lie outside of the lipid bilayer, both extracellular or intracellular.

The present invention contemplates the isolation of additional closely related species variants. Southern and Northern blot analysis should establish that similar genetic entities exist in other mammals. It is likely that 63954s are widespread in species variants, e.g., rodents, lagomorphs, carnivores, artiodactyla, perissodactyla, and primates.

The invention also provides means to isolate a group of related antigens displaying both distinctness and similarities in structure, expression, and function. Elucidation of many of the physiological effects of the molecules will be greatly accelerated by the isolation and characterization of additional distinct species variants of them. In particular, the present invention provides useful probes for identifying additional homologous genetic entities in different species.

The isolated genes will allow transformation of cells lacking expression of a corresponding 63954, e.g., either species types or cells which lack corresponding antigens and

exhibit negative background activity. This should allow analysis of the function of 63954 in comparison to untransformed control cells.

Dissection of critical structural elements which effect the various activation or differentiation functions mediated through these antigens is possible using standard techniques of modern molecular biology, particularly in comparing members of the related class. See, e.g., the homolog-scanning mutagenesis technique described in Cunningham, et al. (1989) Science 243:1339-1336; and approaches used in O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992; and Lechleiter, et al. (1990) EMBO J. 9:4381-4390.

Intracellular functions would probably involve segments of the antigen which are normally accessible to the cytosol. However, protein internalization may occur under certain circumstances, and interaction between intracellular components and "extracellular" segments may occur. The specific segments of interaction of 63954 with other intracellular components may be identified by mutagenesis or direct biochemical means, e.g., cross-linking or affinity methods. Structural analysis by crystallographic or other physical methods will also be applicable. Further investigation of the mechanism of signal transduction will include study of associated components which may be isolatable by affinity methods or by genetic means, e.g., complementation analysis of mutants.

Further study of the expression and control of 63954 will be pursued. The controlling elements associated with the antigens should exhibit differential physiological, developmental, tissue specific, or other expression patterns. Upstream or downstream genetic regions, e.g., control elements, are of interest. In particular, physiological or developmental variants, e.g., multiple alternatively processed forms of the antigen might be found. See, e.g., SEQ ID NO: 1, 3, and 5. Thus, differential splicing of message may lead to an assortment of membrane

bound forms, soluble forms, and modified versions of antigen.

Structural studies of the antigens will lead to design of new antigens, particularly analogs exhibiting agonist or antagonist properties on the molecule. This can be combined with previously described screening methods to isolate antigens exhibiting desired spectra of activities.

V. Antibodies

Antibodies can be raised to various 63954s, including species, polymorphic, or allelic variants, and fragments thereof, both in their naturally occurring forms and in their recombinant forms. Additionally, antibodies can be raised to 63954s in either their active forms or in their inactive forms, including native or denatured versions. Anti-idiotypic antibodies are also contemplated.

Antibodies, including binding fragments and single chain versions, against predetermined fragments of the antigens can be raised by immunization of animals with conjugates of the fragments with immunogenic proteins. Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies can be screened for binding to normal or defective 63954s, or screened for agonistic or antagonistic activity, e.g., mediated through the antigen or its binding partner. Antibodies may be agonistic or antagonistic, e.g., by sterically blocking ligand binding. These monoclonal antibodies will usually bind with at least a K_D of about 1 mM, more usually at least about 300 µM, typically at least about 100 µM, more typically at least about 30 µM, preferably at least about 10 µM, and more preferably at least about 3 µM or better.

The antibodies of this invention can also be useful in diagnostic applications. As capture or non-neutralizing antibodies, they can be screened for ability to bind to the antigens without inhibiting binding by a partner. As neutralizing antibodies, they can be useful in competitive binding assays. They will also be useful in detecting or

quantifying 63954 protein or its 28 binding partners. See, e.g., Chan (ed.) (1987) Immunology: A Practical Guide, Academic Press, Orlando, FL; Price and Newman (eds.) (1991) Principles and Practice of Immunoassay, Stockton Press, N.Y.; and Ngo (ed.) (1988) Nonisotopic Immunoassay, Plenum Press, N.Y. Cross absorptions or other tests will identify antibodies which exhibit various spectra of specificities, e.g., unique or shared species specificities.

Further, the antibodies, including antigen binding fragments, of this invention can be potent antagonists that bind to the antigen and inhibit functional binding or inhibit the ability of a binding partner to elicit a biological response. They also can be useful as non-neutralizing antibodies and can be coupled to toxins or radionuclides so that when the antibody binds to antigen, a cell expressing it, e.g., on its surface, is killed. Further, these antibodies can be conjugated to drugs or other therapeutic agents, either directly or indirectly by means of a linker, and may effect drug targeting.

Antigen fragments may be joined to other materials, particularly polypeptides, as fused or covalently joined polypeptides to be used as immunogens. An antigen and its fragments may be fused or covalently linked to a variety of immunogens, such as keyhole limpet hemocyanin, bovine serum albumin, tetanus toxoid, etc. See Microbiology, Hoeber Medical Division, Harper and Row, 1969; Landsteiner (1962) Specificity of Serological Reactions, Dover Publications, New York; Williams, et al. (1967) Methods in Immunology and Immunochemistry, vol. 1, Academic Press, New York; and Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press, NY, for descriptions of methods of preparing polyclonal antisera.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of techniques for preparing such monoclonal antibodies may be found in, e.g., Stites, et al. (eds.) Basic and Clinical

Immunology (4th ed.), Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.), Academic Press, New York; and particularly in Kohler and Milstein (1975) in Nature 256:495-497, which discusses one method of generating monoclonal antibodies.

Other suitable techniques involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors. See, Huse, et al. (1989) "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda," Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546. The polypeptides and antibodies of the present invention may be used with or without modification, including chimeric or humanized antibodies. Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents, teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced, see Cabilly, U.S. Patent No. 4,816,567; Moore, et al., U.S. Patent No. 4,642,334; and Queen, et al. (1989) Proc. Nat'l Acad. Sci. USA 86:10029-10033.

The antibodies of this invention can also be used for affinity chromatography in isolating the protein. Columns can be prepared where the antibodies are linked to a solid support. See, e.g., Wilchek et al. (1984) Meth. Enzymol. 104:3-55.

Antibodies raised against each 63954 will also be useful to raise anti-idiotypic antibodies. These will be

useful in detecting or diagnosing various immunological conditions related to expression of the respective antigens.

VI. Nucleic Acids

The described peptide sequences and the related reagents are useful in detecting, isolating, or identifying a DNA clone encoding 63954, e.g., from a natural source. Typically, it will be useful in isolating a gene from mammal, and similar procedures will be applied to isolate genes from other species, e.g., warm blooded animals, such as birds and mammals. Cross hybridization will allow isolation of 63954 from other species. A number of different approaches should be available to successfully isolate a suitable nucleic acid clone.

The purified protein or defined peptides are useful for generating antibodies by standard methods, as described above. Synthetic peptides or purified protein can be presented to an immune system to generate monoclonal or polyclonal antibodies. See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene; and Harlow and Lane (1989) Antibodies: A Laboratory Manual, Cold Spring Harbor Press. Alternatively, the 63954 can be used as a specific binding reagent, and advantage can be taken of its specificity of binding, much like an antibody would be used.

For example, the specific binding composition could be used for screening of an expression library made from a cell line which expresses a 63954. The screening can be standard staining of surface expressed antigen, or by panning. Screening of intracellular expression can also be performed by various staining or immunofluorescence procedures. The binding compositions could be used to affinity purify or sort out cells expressing the protein.

The peptide segments can also be used to predict appropriate oligonucleotides to screen a library. The genetic code can be used to select appropriate oligonucleotides useful as probes for screening. See, e.g., SEQ ID NO: 1, 3, 5, and 7. In combination with polymerase

chain reaction (PCR) techniques, synthetic oligonucleotides will be useful in selecting correct clones from a library. Complementary sequences will also be used as probes, primers, or antisense strands. Based upon identification of the likely extracellular domain, various fragments should be particularly useful, e.g., coupled with anchored vector or poly-A complementary PCR techniques or with complementary DNA of other peptides.

This invention contemplates use of isolated DNA or fragments to encode a biologically active corresponding 63954 polypeptide. In addition, this invention covers isolated or recombinant DNA which encodes a biologically active protein or polypeptide which is capable of hybridizing under appropriate conditions with the DNA sequences described herein. Said biologically active protein or polypeptide can be an intact antigen, or fragment, and have an amino acid sequence disclosed in, e.g., SEQ ID NO: 2, 4, 6, or 8. Further, this invention covers the use of isolated or recombinant DNA, or fragments thereof, which encode proteins which are homologous to a 63954 or which was isolated using cDNA encoding a 63954 as a probe. The isolated DNA can have the respective regulatory sequences in the 5' and 3' flanks, e.g., promoters, enhancers, poly-A addition signals, and others.

An "isolated" nucleic acid is a nucleic acid, e.g., an RNA, DNA, or a mixed polymer, which is substantially separated from other components which naturally accompany a native sequence, e.g., ribosomes, polymerases, and/or flanking genomic sequences from the originating species. The term embraces a nucleic acid sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems. A substantially pure molecule includes isolated forms of the molecule. Generally, the nucleic acid will be in a vector or fragment less than about

50 kb, usually less than about 30 kb, typically less than about 10 kb, and preferably less than about 6 kb.

An isolated nucleic acid will generally be a homogeneous composition of molecules, but will, in some embodiments, contain minor heterogeneity. This heterogeneity is typically found at the polymer ends or portions not critical to a desired biological function or activity.

A "recombinant" nucleic acid is defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence, typically selection or production. Alternatively, it can be a nucleic acid made by generating a sequence comprising fusion of two fragments which are not naturally contiguous to each other, but is meant to exclude products of nature, e.g., naturally occurring mutants. Thus, e.g., products made by transforming cells with any unnaturally occurring vector is encompassed, as are nucleic acids comprising sequence derived using any synthetic oligonucleotide process. Such is often done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site.

Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a single genetic entity comprising a desired combination of functions not found in the commonly available natural forms. Restriction enzyme recognition sites are often the target of such artificial manipulations, but other site specific targets, e.g., promoters, DNA replication sites, regulation sequences, control sequences, or other useful features may be incorporated by design. A similar concept is intended for a recombinant, e.g., fusion, polypeptide. Specifically included are synthetic nucleic acids which, by genetic code redundancy, encode polypeptides similar to fragments of

these antigens, and fusions of sequences from various different species variants.

A significant "fragment" in a nucleic acid context is a contiguous segment of at least about 17 nucleotides, generally at least about 22 nucleotides, ordinarily at least about 29 nucleotides, more often at least about 35 nucleotides, typically at least about 41 nucleotides, usually at least about 47 nucleotides, preferably at least about 55 nucleotides, and in particularly preferred embodiments will be at least about 60 or more nucleotides.

A DNA which codes for a 63954 protein will be particularly useful to identify genes, mRNA, and cDNA species which code for related or homologous proteins, as well as DNAs which code for homologous proteins from different species. There are likely homologs in other species, including primates, rodents, and birds. Various 63954 proteins should be homologous and are encompassed herein. However, even genes encoding proteins that have a more distant evolutionary relationship to the antigen can readily be isolated under appropriate conditions using these sequences if they are sufficiently homologous. Primate 63954 proteins are of particular interest.

Recombinant clones derived from the genomic sequences, e.g., containing introns, will be useful for transgenic studies, including, e.g., transgenic cells and organisms, and for gene therapy. See, e.g., Goodnow (1992) "Transgenic Animals" in Roitt (ed.) Encyclopedia of Immunology, Academic Press, San Diego, pp. 1502-1504; Travis (1992) Science 256:1392-1394; Kuhn, et al. (1991) Science 254:707-710; Capecchi (1989) Science 244:1288; Robertson (1987)(ed.) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Oxford; and Rosenberg (1992) J. Clinical Oncology 10:180-199.

Substantial homology in the nucleic acid sequence comparison context means either that the segments, or their complementary strands, when compared, are identical when optimally aligned, with appropriate nucleotide insertions or

deletions, in at least about 50% of the nucleotides, generally at least about 58%, ordinarily at least about 65%, often at least about 71%, typically at least about 77%, usually at least about 85%, preferably at least about 95 to 98% or more, and in particular embodiments, as high as about 99% or more of the nucleotides. Alternatively, substantial homology exists when the segments will hybridize under selective hybridization conditions, to a strand, or its complement, typically using a sequence of 63954, e.g., in SEQ ID NO: 1, 3, 5, or 7. Typically, selective hybridization will occur when there is at least about 55% homology over a stretch of at least about 30 nucleotides, preferably at least about 75% over a stretch of about 25 nucleotides, and most preferably at least about 90% over about 20 nucleotides. See, Kanehisa (1984) Nuc. Acids Res. 12:203-213. The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, usually at least about 28 nucleotides, typically at least about 40 nucleotides, and preferably at least about 75 to 100 or more nucleotides.

Stringent conditions, in referring to homology in the hybridization context, will be stringent combined conditions of salt, temperature, organic solvents, and other parameters, typically those controlled in hybridization reactions. Stringent temperature conditions will usually include temperatures in excess of about 30° C, usually in excess of about 37° C, typically in excess of about 55° C, preferably in excess of about 70° C. Stringent salt conditions will ordinarily be less than about 1000 mM, usually less than about 400 mM, typically less than about 250 mM, preferably less than about 150 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Wetmur and Davidson (1968) J. Mol. Biol. 31:349-370. 63954 from other mammalian species can be cloned and isolated by cross-species hybridization of closely related species. Homology

may be relatively low between distantly related species, and thus hybridization of relatively closely related species is advisable. Alternatively, preparation of an antibody preparation which exhibits less species specificity may be useful in expression cloning approaches.

VII. Making 63954; Mimetics

DNA which encodes the 63954 or fragments thereof can be obtained by chemical synthesis, screening cDNA libraries, or screening genomic libraries prepared from a wide variety of cell lines or tissue samples. See, e.g., Okayama and Berg (1982) Mol. Cell. Biol. 2:161-170; Gubler and Hoffman (1983) Gene 25:263-269; and Glover (ed.) (1984) DNA Cloning: A Practical Approach, IRL Press, Oxford. Alternatively, the sequences provided herein provide useful PCR primers or allow synthetic or other preparation of suitable genes encoding a 63954; including, naturally occurring embodiments.

This DNA can be expressed in a wide variety of host cells for the synthesis of a full-length 63954 or fragments which can in turn, e.g., be used to generate polyclonal or monoclonal antibodies; for binding studies; for construction and expression of modified molecules; and for structure/function studies.

Vectors, as used herein, comprise plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles which enable the integration of DNA fragments into the genome of the host. See, e.g., Pouwels, et al. (1985 and Supplements) Cloning Vectors: A Laboratory Manual, Elsevier, N.Y.; and Rodriguez, et al. (1988)(eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Buttersworth, Boston, MA.

For purposes of this invention, DNA sequences are operably linked when they are functionally related to each other. For example, DNA for a presequence or secretory leader is operably linked to a polypeptide if it is expressed as a preprotein or participates in directing the

polypeptide to the cell membrane or in secretion of the polypeptide. A promoter is operably linked to a coding sequence if it controls the transcription of the polypeptide; a ribosome binding site is operably linked to a coding sequence if it is positioned to permit translation. Usually, operably linked means contiguous and in reading frame, however, certain genetic elements such as repressor genes are not contiguously linked but still bind to operator sequences that in turn control expression. See e.g., Rodriguez, et al., Chapter 10, pp. 205-236; Balbas and Bolivar (1990) Methods in Enzymol. 185:14-37; and Ausubel, et al. (1993) Current Protocols in Molecular Biology, Greene and Wiley, NY.

Representative examples of suitable expression vectors include pCDNA1; pCD, see Okayama, et al. (1985) Mol. Cell Biol. 5:1136-1142; pMC1neo Poly-A, see Thomas, et al. (1987) Cell 51:503-512; and a baculovirus vector such as pAC 373 or pAC 610. See, e.g., Miller (1988) Ann. Rev. Microbiol. 42:177-199.

It will often be desired to express a 63954 polypeptide in a system which provides a specific or defined glycosylation pattern. See, e.g., Luckow and Summers (1988) Bio/Technology 6:47-55; and Kaufman (1990) Meth. Enzymol. 185:487-511.

The 63954, or a fragment thereof, may be engineered to be phosphatidyl inositol (PI) linked to a cell membrane, but can be removed from membranes by treatment with a phosphatidyl inositol cleaving enzyme, e.g., phosphatidyl inositol phospholipase-C. This releases the antigen in a biologically active form, and allows purification by standard procedures of protein chemistry. See, e.g., Low (1989) Biochim. Biophys. Acta 988:427-454; Tse, et al. (1985) Science 230:1003-1008; and Brunner, et al. (1991) J. Cell Biol. 114:1275-1283.

Now that the 63954 has been characterized, fragments or derivatives thereof can be prepared by conventional processes for synthesizing peptides. These include

processes such as are described in Stewart and Young (1984) Solid Phase Peptide Synthesis, Pierce Chemical Co., Rockford, IL; Bodanszky and Bodanszky (1984) The Practice of Peptide Synthesis, Springer-Verlag, New York, NY; Bodanszky (1984) The Principles of Peptide Synthesis, Springer-Verlag, New York; and Villafranca (ed.) (1991) Techniques in Protein Chemistry II, Academic Press, San Diego, Ca.

VIII. Uses

The present invention provides reagents which will find use in diagnostic applications as described elsewhere herein, e.g., in the general description for T cell mediated conditions, or below in the description of kits for diagnosis.

This invention also provides reagents with significant therapeutic value. The 63954 (naturally occurring or recombinant), fragments thereof, and antibodies thereto, along with compounds identified as having binding affinity to 63954, should be useful in the treatment of conditions associated with abnormal physiology or development, including abnormal proliferation, e.g., cancerous conditions, or degenerative conditions. In particular, modulation of development of lymphoid cells will be achieved by appropriate therapeutic treatment using the compositions provided herein. For example, a disease or disorder associated with abnormal expression or abnormal signaling by a 63954 should be a likely target for an agonist or antagonist of the antigen. The antigen plays a role in regulation or development of hematopoietic cells, e.g., lymphoid cells, which affect immunological responses, e.g., autoimmune disorders.

In particular, the antigen will likely provide a costimulatory signal to cell activation. Thus, the 63954 will likely modulate T cell mediated interactions with other cell types, e.g., cells which possess a receptor therefor. These interactions would lead, in particular contexts, to

modulation of cell growth, cytokine synthesis by those or other cells, or development of particular effector cells.

Moreover, the 63954 or antagonists could redirect T cell responses, e.g., between Th1 and Th2 polarization, or with Th0 cells. Among these agonists should be various antibodies which recognize the appropriate epitopes, e.g., which mimic binding of 63954 to its receptor. Alternatively, they may bind to epitopes which sterically can block receptor binding.

Antagonists of 63954, such as the naturally occurring secreted form of 63954 or blocking antibodies, may also be useful, They may provide a selective and powerful way to modulate immune responses in abnormal situations, e.g., autoimmune disorders, including rheumatoid arthritis, systemic lupus erythematosis (SLE), Hashimoto's autoimmune thyroiditis, as well as acute and chronic inflammatory responses in which T cell activation, expansion, and/or immunological T cell memory play an important role. See also Samter, et al. (eds) Immunological Diseases vols. 1 and 2, Little, Brown and Co. Regulation of T cell activation, expansion, and/or cytokine release by the naturally occurring secreted form of 63954, or an antagonist thereof, may be effected.

In addition, certain combination compositions with other modulators of T cell signaling would be useful. Such other signaling molecules include TCR reagents, CD40, CD40L, CTLA-8, CD28, SLAM, FAS, and their respective antagonists.

Various abnormal conditions are known in each of the cell types shown to possess 63954 mRNA by Northern blot analysis. See Berkow (ed.) The Merck Manual of Diagnosis and Therapy, Merck & Co., Rahway, NJ; Thorn, et al. Harrison's Principles of Internal Medicine, McGraw-Hill, NY; and Weatherall, et al. (eds.) Oxford Textbook of Medicine, Oxford University Press, Oxford. Many other medical conditions and diseases involve T cells or are T cell mediated, and many of these will be responsive to treatment by an agonist or antagonist provided herein. See, e.g.,

Stites and Terr (eds; 1991) Basic and Clinical Immunology Appleton and Lange, Norwalk, CT; and Samter, et al. (eds) Immunological Diseases Little, Brown and Co. These problems should be susceptible to prevention or treatment using compositions provided herein.

63954 antibodies can be purified and then administered to a patient, veterinary or human. These reagents can be combined for therapeutic use with additional active or inert ingredients, e.g., in conventional pharmaceutically acceptable carriers or diluents, e.g., immunogenic adjuvants, along with physiologically innocuous stabilizers, excipients, or preservatives. These combinations can be sterile filtered and placed into dosage forms as by lyophilization in dosage vials or storage in stabilized aqueous preparations. This invention also contemplates use of antibodies or binding fragments thereof, including forms which are not complement binding.

Drug screening using 63954 or fragments thereof can be performed to identify compounds having binding affinity to or other relevant biological effects on 63954 functions, including isolation of associated components. Subsequent biological assays can then be utilized to determine if the compound has intrinsic stimulating activity or is a blocker or antagonist in that it blocks the activity of the antigen, e.g., mutein antagonists. Likewise, a compound having intrinsic stimulating activity can activate the signal pathway and is thus an agonist in that it simulates the activity of 63954. This invention further contemplates the therapeutic use of blocking antibodies to 63954 as antagonists and of stimulatory molecules, e.g., muteins, as agonists. This approach should be particularly useful with other 63954 species variants.

The quantities of reagents necessary for effective therapy will depend upon many different factors, including means of administration, target site, physiological state of the patient, and other medicants administered. Thus, treatment dosages should be titrated to optimize safety and

efficacy. Typically, dosages used in vitro may provide useful guidance in the amounts useful for in situ administration of these reagents. Animal testing of effective doses for treatment of particular disorders will provide further predictive indication of human dosage. Various considerations are described, e.g., in Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn. Methods for administration are discussed therein and below, e.g., for oral, intravenous, intraperitoneal, or intramuscular administration, transdermal diffusion, and others. Pharmaceutically acceptable carriers will include water, saline, buffers, and other compounds described, e.g., in the Merck Index, Merck & Co., Rahway, New Jersey. Dosage ranges would ordinarily be expected to be in amounts lower than 1 mM concentrations, typically less than about 10 μM concentrations, usually less than about 100 nM, preferably less than about 10 pM (picomolar), and most preferably less than about 1 fM (femtomolar), with an appropriate carrier. Slow release formulations, or a slow release apparatus will often be utilized for continuous or long term administration. See, e.g., Langer (1990) Science 249:1527-1533.

63954, fragments thereof, and antibodies to it or its fragments, antagonists, and agonists, may be administered directly to the host to be treated or, depending on the size of the compounds, it may be desirable to conjugate them to carrier proteins such as ovalbumin or serum albumin prior to their administration. Therapeutic formulations may be administered in many conventional dosage formulations. While it is possible for the active ingredient to be administered alone, it is preferable to present it as a pharmaceutical formulation. Formulations typically comprise at least one active ingredient, as defined above, together with one or more acceptable carriers thereof. Each carrier

should be both pharmaceutically and physiologically acceptable in the sense of being compatible with the other ingredients and not injurious to the patient. Formulations include those suitable for oral, rectal, nasal, topical, or parenteral (including subcutaneous, intramuscular, intravenous and intradermal) administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. See, e.g., Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn.; Avis, et al. (eds.) (1993) Pharmaceutical Dosage Forms: Parenteral Medications, Dekker, New York; Lieberman, et al. (eds.) (1990) Pharmaceutical Dosage Forms: Tablets, Dekker, New York; and Lieberman, et al. (eds.) (1990) Pharmaceutical Dosage Forms: Disperse Systems, Dekker, New York. The therapy of this invention may be combined with or used in association with other agents, e.g., other modulators of T cell activation, e.g., CD40, CD40 ligand, CD28, CTLA-4, B7, B70, SLAM, T cell receptor signaling entities, or their respective antagonists.

Both the naturally occurring and the recombinant form of the 63954s of this invention are particularly useful in kits and assay methods which are capable of screening compounds for binding activity to the proteins. Several methods of automating assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period. See, e.g., Fodor, et al. (1991) Science 251:767-773, which describes means for testing of binding affinity by a plurality of defined polymers synthesized on a solid substrate. The development of suitable assays can be greatly facilitated by the availability of large amounts of purified, soluble 63954 as provided by this invention.

Other methods can be used to determine the critical residues in the 63954-63954 receptor interactions. Mutational analysis can be performed, e.g., see Somoza, et al. (1993) J. Exp. Med. 178:549-558, to determine specific residues critical in the interaction and/or signaling. Both extracellular domains, involved in the homophilic interaction, or intracellular domain, which provides interactions important in intracellular signaling.

For example, antagonists can normally be found once the antigen has been structurally defined, e.g., by tertiary structure data. Testing of potential interacting analogs is now possible upon the development of highly automated assay methods using a purified 63954. In particular, new agonists and antagonists will be discovered by using screening techniques described herein. Of particular importance are compounds found to have a combined binding affinity for a spectrum of 63954 molecules, e.g., compounds which can serve as antagonists for species variants of 63954.

One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant DNA molecules expressing a 63954. Cells may be isolated which express a 63954 in isolation from other molecules. Such cells, either in viable or fixed form, can be used for standard binding partner binding assays. See also, Parce, et al. (1989) Science 246:243-247; and Owicki, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:4007-4011, which describe sensitive methods to detect cellular responses.

Another technique for drug screening involves an approach which provides high throughput screening for compounds having suitable binding affinity to a 63954 and is described in detail in Geysen, European Patent Application 84/03564, published on September 13, 1984. First, large numbers of different small peptide test compounds are synthesized on a solid substrate, e.g., plastic pins or some other appropriate surface, see Fodor, et al. (1991). Then all the pins are reacted with solubilized, unpurified or

solubilized, purified 63954, and washed. The next step involves detecting bound 63954.

Rational drug design may also be based upon structural studies of the molecular shapes of the 63954 and other effectors or analogs. Effectors may be other proteins which mediate other functions in response to binding, or other proteins which normally interact with 63954. One means for determining which sites interact with specific other proteins is a physical structure determination, e.g., x-ray crystallography or 2 dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, e.g., Blundell and Johnson (1976) Protein Crystallography, Academic Press, New York.

IX. Kits

This invention also contemplates use of 63954 proteins, fragments thereof, peptides, and their fusion products in a variety of diagnostic kits and methods for detecting the presence of another 63954 or binding partner. Typically the kit will have a compartment containing either a defined 63954 peptide or gene segment or a reagent which recognizes one or the other, e.g., 63954 fragments or antibodies.

A kit for determining the binding affinity of a test compound to a 63954 would typically comprise a test compound; a labeled compound, for example a binding partner or antibody having known binding affinity for 63954; a source of 63954 (naturally occurring or recombinant); and a means for separating bound from free labeled compound, such as a solid phase for immobilizing the molecule. Once compounds are screened, those having suitable binding affinity to the antigen can be evaluated in suitable biological assays, as are well known in the art, to determine whether they act as agonists or antagonists to the 63954 signaling pathway. The availability of recombinant

63954 polypeptides also provide well defined standards for calibrating such assays.

A preferred kit for determining the concentration of, e.g., a 63954 in a sample would typically comprise a labeled compound, e.g., binding partner or antibody, having known binding affinity for the antigen, a source of antigen (naturally occurring or recombinant) and a means for separating the bound from free labeled compound, e.g., a solid phase for immobilizing the 63954. Compartments containing reagents, and instructions, will normally be provided.

Antibodies, including antigen binding fragments, specific for the 63954 or fragments are useful in diagnostic applications to detect the presence of elevated levels of 63954 and/or its fragments. Such diagnostic assays can employ lysates, live cells, fixed cells, immunofluorescence, cell cultures, body fluids, and further can involve the detection of antigens related to the antigen in serum, or the like. Diagnostic assays may be homogeneous (without a separation step between free reagent and antigen-binding partner complex) or heterogeneous (with a separation step). Various commercial assays exist, such as radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), enzyme immunoassay (EIA), enzyme-multiplied immunoassay technique (EMIT), substrate-labeled fluorescent immunoassay (SLFIA), and the like. See, e.g., Van Vunakis, et al. (1980) Meth Enzymol. 70:1-525; Harlow and Lane (1980) Antibodies: A Laboratory Manual, CSH Press, NY; and Coligan, et al. (eds.) (1993) Current Protocols in Immunology, Greene and Wiley, NY.

Anti-idiotypic antibodies may have similar use to diagnose presence of antibodies against a 63954, as such may be diagnostic of various abnormal states. For example, overproduction of 63954 may result in production of various immunological reactions which may be diagnostic of abnormal physiological states, particularly in proliferative cell

conditions such as cancer or abnormal activation or differentiation.

Frequently, the reagents for diagnostic assays are supplied in kits, so as to optimize the sensitivity of the assay. For the subject invention, depending upon the nature of the assay, the protocol, and the label, either labeled or unlabeled antibody or binding partner, or labeled 63954 is provided. This is usually in conjunction with other additives, such as buffers, stabilizers, materials necessary for signal production such as substrates for enzymes, and the like. Preferably, the kit will also contain instructions for proper use and disposal of the contents after use. Typically the kit has compartments for each useful reagent. Desirably, the reagents are provided as a dry lyophilized powder, where the reagents may be reconstituted in an aqueous medium providing appropriate concentrations of reagents for performing the assay.

Many of the aforementioned constituents of the drug screening and the diagnostic assays may be used without modification or may be modified in a variety of ways. For example, labeling may be achieved by covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In any of these assays, the binding partner, test compound, 63954, or antibodies thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups: radiolabels such as ^{125}I, enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin coupled to one of the above label groups.

There are also numerous methods of separating the bound from the free 63954, or alternatively the bound from the free test compound. The 63954 can be immobilized on various matrixes followed by washing. Suitable matrixes

include plastic such as an ELISA plate, filters, and beads. See, e.g., Coligan, et al. (eds.) (1993) Current Protocols in Immunology, Vol. 1, Chapter 2, Greene and Wiley, NY. Other suitable separation techniques include, without limitation, the fluorescein antibody magnetizable particle method described in Rattle, et al. (1984) Clin. Chem. 30:1457-1461, and the double antibody magnetic particle separation as described in U.S. Pat. No. 4,659,678.

Methods for linking proteins or their fragments to the various labels have been extensively reported in the literature and do not require detailed discussion here. Many of the techniques involve the use of activated carboxyl groups either through the use of carbodiimide or active esters to form peptide bonds, the formation of thioethers by reaction of a mercapto group with an activated halogen such as chloroacetyl, or an activated olefin such as maleimide, for linkage, or the like. Fusion proteins will also find use in these applications.

Another diagnostic aspect of this invention involves use of oligonucleotide or polynucleotide sequences taken from the sequence of a 63954. These sequences can be used as probes for detecting levels of the 63954 message in samples from patients suspected of having an abnormal condition, e.g., cancer or developmental problem. Since the antigen is a marker for activation, it may be useful to determine the numbers of activated T cells to determine, e.g., when additional suppression may be called for. The preparation of both RNA and DNA nucleotide sequences, the labeling of the sequences, and the preferred size of the sequences has received ample description and discussion in the literature. See, e.g., Langer-Safer, et al. (1982) Proc. Nat'l. Acad. Sci. 79:4381-4385; Caskey (1987) Science 236:962-967; and Wilchek et al. (1988) Anal. Biochem. 171:1-32.

Diagnostic kits which also test for the qualitative or quantitative presence of other markers are also contemplated. Diagnosis or prognosis may depend on the

combination of multiple indications used as markers. Thus, kits may test for combinations of markers. See, e.g., Viallet, et al. (1989) Progress in Growth Factor Res. 1:89-97. Other kits may be used to evaluate T cell subsets.

X. Methods for Isolating 63954 Specific Binding Partners

The 63954 protein should interact with a receptor based, e.g., upon its similarity in structure and function to other cell surface antigens exhibiting similar structure and cell type specificity of expression. Methods to isolate a receptor are made available by the ability to make purified 63954 for screening programs. Soluble or other constructs using the 63954 sequences provided herein will allow for screening or isolation of 63954 specific receptors. Many methods exist for expression cloning, panning, affinity isolation, or other means to identify a receptor. A two-hybrid selection system may also be applied making appropriate constructs with the available 63954 sequences. See, e.g., Fields and Song (1989) Nature 340:245-246.

EXAMPLES

General Methods

Some of the standard methods are described or referenced, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.), vols. 1-3, CSH Press, NY; Ausubel, et al., Biology, Greene Publishing Associates, Brooklyn, NY; or Ausubel, et al. (1987 and Supplements) Current Protocols in Molecular Biology, Greene and Wiley, New York; Innis, et al. (eds.)(1990) PCR Protocols: A Guide to Methods and Applications, Academic Press, N.Y. Methods for protein purification include such methods as ammonium sulfate precipitation, column chromatography, electrophoresis, centrifugation, crystallization, and others. See, e.g., Ausubel, et al. (1987 and periodic

supplements); Deutscher (1990) "Guide to Protein Purification" in Methods in Enzymol. vol. 182, and other volumes in this series; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, N.J., or Bio-Rad, Richmond, CA. Combination with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See, e.g., Hochuli (1989) Chemische Industrie 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, N.Y.; and Crowe, et al. (1992) QIAexpress: The High Level Expression & Protein Purification System QIAGEN, Inc., Chatsworth, CA. Cell culture techniques are described in Doyle, et al. (eds.) (1994) Cell and Tissue Culture: Laboratory Procedures, John Wiley and Sons, NY.

Standard immunological techniques are described, e.g., in Hertzenberg, et al. (eds. 1996) Weir's Handbook of Experimental Immunology vols. 1-4, Blackwell Science; Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; and Methods in Enzymology volumes. 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162, and 163.

FACS analyses are described in Melamed, et al. (1990) Flow Cytometry and Sorting Wiley-Liss, Inc., New York, NY; Shapiro (1988) Practical Flow Cytometry Liss, New York, NY; and Robinson, et al. (1993) Handbook of Flow Cytometry Methods Wiley-Liss, New York, NY. Fluorescent labeling of appropriate reagents was performed by standard methods.

EXAMPLE 1: Cloning of Human 63954

63954 was assembled by careful analysis of ESTs found in various databases. These ESTs were from cDNA libraries derived from human B cell lymphoma, human bone marrow, PHA activated T-cells, CD34 depleted cord blood, primary dendritic cells, human T-cell lymphoma, macrophage-oxLDL, bone marrow cell line RS4;11, human neutrophil, smooth

muscle, stomach cancer, and Soares fetal liver. PCR primers are designed and synthesized and a PCR product is obtained from any of these libraries. This product is used as a hybridization clone to screen these libraries for a full length clone.

EXAMPLE 2: Cellular Expression of Human 63954

A probe specific for cDNA encoding human 63954 is used to determine tissue distribution of message encoding the antigen. Standard hybridization probes may be used to do a Northern analysis of RNA from appropriate sources, either cells, e.g., stimulated or in various physiological states, in various tissues, e.g., spleen, liver, thymus, lung, etc., or in various species. Southern analysis of cDNA libraries may also provide valuable distribution information. Standard tissue blots or species blots are commercially available. Similar techniques will be useful for evaluating diagnostic or medical conditions which may correlate with expression in various cell types.

PCR analysis using appropriate primers may also be used. Antibody analysis, including immunohistochemistry or FACS, may be used to determine cellular or tissue distribution.

Southern blot analysis of cDNA libraries were performed on: U937 premonocytic line, resting (M100); elutriated monocytes, activated with LPS, IFNγ, anti-IL-10 for 4, 16 h pooled (M106); elutriated monocytes, activated with LPS, IFNγ, IL-10 for 4, 16 h pooled (M107); elutriated monocytes, activated LPS for 1 h (M108); elutriated monocytes, activated LPS for 6 h (M109); dendritic cells (DC) 30% CD1a+, from CD34+ GM-CSF, TNFα 12 days, resting; DC 70% CD1a+, from CD34+ GM-CSF, TNFα 12 days, resting (D101); DC 70% CD1a+, from CD34+ GM-CSF, TNFα 12 days, activated with PMA and ionomycin for 1 hr (D102); DC 70% CD1a+, from CD34+ GM-CSF, TNFα 12 days, activated with PMA and ionomycin for 6 hr (D103); DC 95% CD1a+, from CD34+ GM-CSF, TNFα 12 days activated with PMA and ionomycin for 1 or

6 hr, pooled; DC from monocytes GM-CSF, IL-4 5 days, resting (D107); DC from monocytes GM-CSF, IL-4 5 days, resting (D108); DC from monocytes GM-CSF, IL-4 5 days, activated TNFα, monocyte supe for 4, 16 h pooled (D110); EBV tranfected B cell lines, resting; spleenocytes, resting; spleenocytes, activated with PMA and ionomycin; 20 NK clones resting, pooled; 20 NK clones activated with PMA and ionomycin, pooled; NKL clone, IL-2 treated; NK cytotoxic clone, resting; adipose tissue fetal 28 wk male (O108); brain fetal 28 wk male (O104); gallbladder fetal 28 wk male (O106); heart fetal 28 wk male (O103); small intestine fetal 28 wk male (O107); kidney fetal 28 wk male (O100); liver fetal 28 wk male (O102); lung fetal 28 wk male (O101); ovary fetal 25 wk female (O109); adult placenta 28 wk (O113); spleen fetal 28 wk male (O112); testes fetal 28 wk male (O111); uterus fetal 25 wk female (O110); TH0 clone Mot 72, resting (T102); T cell, TH0 clone Mot 72, activated with anti-CD28 and anti-CD3 for 3, 6, 12 h pooled (T103); T cell, TH0 clone Mot 72, anergic treated with specific peptide for 2, 7, 12 h pooled (T104); Th0 subtraction of resting from activated; T cell, TH1 clone HY06, resting (T107); T cell, TH1 clone HY06, activated with anti-CD28 and anti-CD3 for 3, 6, 12 h pooled (T108); T cell, TH1 clone HY06, anergic treated with specific peptide for 2, 6, 12 h pooled (T109); Th1 subtraction of resting from activated; T cell, TH2 clone HY935, resting (T110); T cell, TH2 clone HY935, activated with anti-CD28 and anti-CD3 for 2, 7, 12 h pooled (T111); and Th2 subtraction of resting from activated.

A high signal was detected in U937 premonocytic line, resting (M100). Intermediate signals were detected in elutriated monocytes, activated with LPS, IFNγ, anti-IL-10 for 4, 16 h pooled (M106); elutriated monocytes, activated with LPS, IFNγ, IL-10 for 4, 16 h pooled (M107); elutriated monocytes, activated LPS for 1 h (M108); DC 70% CD1a+, from CD34+ GM-CSF, TNFα 12 days, activated with PMA and ionomycin for 6 hr (D103); DC 95% CD1a+, from CD34+ GM-CSF, TNFα 12 days activated with PMA and ionomycin for 1 or 6

51

hr, pooled; and DC from monocytes GM-CSF, IL-4 5 days, resting (D107). Weaker signals were detected in elutriated monocytes, activated LPS for 6 h (M109); DC 70% CD1a+, from CD34+ GM-CSF, TNFα 12 days, resting (D101); DC from monocytes GM-CSF, IL-4 5 days, resting (D108); and DC from monocytes GM-CSF, IL-4 5 days, activated TNFα, monocyte supe for 4, 16 h pooled (D110). No detectable signal was detected in the others.

EXAMPLE 3: Purification of 63954 Protein

Multiple transfected cell lines are screened for one which expresses the antigen, membrane bound or soluble forms, at a high level compared with other cells. Various cell lines are screened and selected for their favorable properties in handling. Natural 63954 can be isolated from natural sources, or by expression from a transformed cell using an appropriate expression vector. Purification of the expressed protein is achieved by standard procedures, or may be combined with engineered means for effective purification at high efficiency from cell lysates or supernatants. FLAG or His_6 segments can be used for such purification features.

EXAMPLE 4: Isolation of Homologous 63954 Genes

The human 63954 cDNA can be used as a hybridization probe to screen a library from a desired source, e.g., a primate cell cDNA library. Many different species can be screened both for stringency necessary for easy hybridization, and for presence using a probe. Appropriate hybridization conditions will be used to select for clones exhibiting specificity of cross hybridization.

Screening by hybridization or PCR using degenerate probes based upon the peptide sequences will also allow isolation of appropriate clones. Alternatively, use of appropriate primers for PCR screening will yield enrichment of appropriate nucleic acid clones.

Similar methods are applicable to isolate either species, polymorphic, or allelic variants. Species variants

are isolated using cross-species hybridization techniques based upon isolation of a full length isolate or fragment from one species as a probe.

Alternatively, antibodies raised against human 63954 will be used to screen for cells which express cross-reactive proteins from an appropriate, e.g., cDNA library. The purified protein or defined peptides are useful for generating antibodies by standard methods, as described above. Synthetic peptides or purified protein are presented to an immune system to generate monoclonal or polyclonal antibodies. See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press. The resulting antibodies are used, e.g., for screening, panning, or sorting.

EXAMPLE 5: Preparation of antibodies specific for 63954

Synthetic peptides or purified protein are presented to an immune system to generate monoclonal or polyclonal antibodies. See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press. Polyclonal serum, or hybridomas may be prepared. In appropriate situations, the binding reagent is either labeled as described above, e.g., fluorescence or otherwise, or immobilized to a substrate for panning methods.

EXAMPLE 6: Isolation of a Receptor for 63954

A 63954 construct expression product can be used as a specific binding reagent to identify its binding partner, e.g., receptor, by taking advantage of its specificity of binding, much like an antibody would be used. A 63954 reagent is either labeled as described above, e.g., fluorescence or otherwise, or immobilized to a substrate for panning methods. See also Anderson, et al. (1997) Nature 390:175-179, which is incorporated herein by reference.

The binding composition is used to screen an expression library made from a cell line which expresses a binding partner, i.e. receptor. Standard staining techniques are used to detect or sort intracellular or surface expressed receptor, or surface expressing transformed cells are screened by panning. Screening of intracellular expression is performed by various staining or immunofluorescence procedures. See also McMahan, et al. (1991) EMBO J. 10:2821-2832.

Alternatively, 63954 reagents are used to affinity purify or sort out cells expressing a receptor. See, e.g., Sambrook, et al. or Ausubel, et al.

Another strategy is to screen for a membrane bound receptor by panning. The cDNA containing receptor cDNA is constructed as described above. The ligand can be immobilized and used to immobilize expressing cells. Immobilization may be achieved by use of appropriate antibodies which recognize, e.g., a FLAG sequence of a 63954 fusion construct, or by use of antibodies raised against the first antibodies. Recursive cycles of selection and amplification lead to enrichment of appropriate clones and eventual isolation of receptor expressing clones.

Phage expression libraries can be screened by 63954. Appropriate label techniques, e.g., anti-FLAG antibodies, will allow specific labeling of appropriate clones.

All citations herein are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Many modifications and variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled.

54

SEQUENCE SUBMISSION

SEQ ID NO: 1 is a primate 63954 nucleic acid sequence.
SEQ ID NO: 2 is a primate 63954 amino acid sequence.
SEQ ID NO: 3 is another primate 63954 nucleic acid sequence.
SEQ ID NO: 4 is another primate 63954 amino acid sequence.
SEQ ID NO: 5 is another primate 63954 nucleic acid sequence.
SEQ ID NO: 6 is another primate 63954 amino acid sequence.
SEQ ID NO: 7 is a rodent 63954 nucleic acid sequence.
SEQ ID NO: 8 is a rodent 63954 amino acid sequence.

(1) GENERAL INFORMATION:

(i) APPLICANT: Schering Corporation.

(ii) TITLE OF INVENTION: Mammalian Cell Surface Antigens; Related Reagents

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Schering-Plough Corporation
(B) STREET: 2000 Galloping Hill Road
(C) CITY: Kenilworth
(D) STATE: New Jersey
(E) COUNTRY: USA
(F) ZIP: 07033-0530

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: Apple Macintosh
(C) OPERATING SYSTEM: Macintosh 7.5.3
(D) SOFTWARE: Microsoft Word 6.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 16-DEC-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/033,601
(B) FILING DATE: 17-DEC-1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Thampoe, Immac J.
(B) REGISTRATION NUMBER: 36,322
(C) REFERENCE/DOCKET NUMBER: DX0688 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (908)298-5061
(B) TELEFAX: (908)298-5388

55

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1109 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 157..1011

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "nucleotide 10 designated C, may be A, C, G, or T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GGGCAGAGAC AAATTCAGGA TAACTCTCCT GAGGGGTGAG CCAAGCCCTG CCATGTAGTG      60

CACGCAGGAC ATCAACAAAC ACAGATAACA GGAAATAATC CATTCCCTGT GGTCACTTAT     120

TCTAAAGGCC CCAACCTTCA AAGTTCAAGT AGTGAT ATG GAT GAC TCC ACA GAA       174
                                        Met Asp Asp Ser Thr Glu
                                          1               5

AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA GAA GAA ATG       222
Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg Glu Glu Met
             10                  15                  20

AAA CTG CAA GGA GTG TGT TTC CAT CCT CCC ACG GAA GGA AAG CCC CTC       270
Lys Leu Gln Gly Val Cys Phe His Pro Pro Thr Glu Gly Lys Pro Leu
         25                  30                  35

TGT CCG ATC CTC CAA AGA CGG AAA GCT GCT GGC TGC AAC CTT GCT GCT       318
Cys Pro Ile Leu Gln Arg Arg Lys Ala Ala Gly Cys Asn Leu Ala Ala
     40                  45                  50

GGC ACT GCT GTC TTG CTG CCT CAC GGT GGT GTC TTT CTA CCA GGT GGC       366
Gly Thr Ala Val Leu Leu Pro His Gly Gly Val Phe Leu Pro Gly Gly
 55                  60                  65                  70

CGC CCT GCA AGG GAC CTG GCC AGC CTC CGG GCA GAG CTG CAG GGC CAC       414
Arg Pro Ala Arg Asp Leu Ala Ser Leu Arg Ala Glu Leu Gln Gly His
                 75                  80                  85
```

```
CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG GCC GGC CTG      462
His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys Ala Gly Leu
             90                  95                 100

GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT GAA CCA CCA      510
Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe Glu Pro Pro
        105                 110                 115

GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT AAG CGT GCC      558
Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn Lys Arg Ala
    120                 125                 130

GTT CAG GGT CCA GAA GAA ACA GTC ACT CAA GAC TGC TTG CAA CTG ATT      606
Val Gln Gly Pro Glu Glu Thr Val Thr Gln Asp Cys Leu Gln Leu Ile
135                 140                 145                 150

GCA GAC AGT GAA ACA CCA ACT ATA CAA AAA GGA TCT TAC ACA TTT GTT      654
Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys Gly Ser Tyr Thr Phe Val
                155                 160                 165

CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA GAA AAA GAG      702
Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu Glu Lys Glu
            170                 175                 180

AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA TAT GGT CAG      750
Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile Tyr Gly Gln
        185                 190                 195

GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA GTT CAG AGG      798
Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu Val Gln Arg
    200                 205                 210

AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG ACT TTG TTT      846
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
215                 220                 225                 230

CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT TCC TGC TAT      894
Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn Ser Cys Tyr
                235                 240                 245

TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC CAA CTT GCA      942
Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu Gln Leu Ala
            250                 255                 260

ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT GTC ACA TTT      990
Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp Val Thr Phe
        265                 270                 275

TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC TGTAGCTATT        1041
Phe Gly Ala Leu Lys Leu Leu
    280                 285

TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC CAAAAAAAAA  1101

AAAAAAAA                                                           1109
```

57

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asp Asp Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu
 1               5                  10                  15

Lys Lys Arg Glu Glu Met Lys Leu Gln Gly Val Cys Phe His Pro Pro
            20                  25                  30

Thr Glu Gly Lys Pro Leu Cys Pro Ile Leu Gln Arg Arg Lys Ala Ala
        35                  40                  45

Gly Cys Asn Leu Ala Ala Gly Thr Ala Val Leu Leu Pro His Gly Gly
    50                  55                  60

Val Phe Leu Pro Gly Gly Arg Pro Ala Arg Asp Leu Ala Ser Leu Arg
65                  70                  75                  80

Ala Glu Leu Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly
                85                  90                  95

Ala Pro Lys Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu
            100                 105                 110

Lys Ile Phe Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn
        115                 120                 125

Ser Arg Asn Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Val Thr Gln
    130                 135                 140

Asp Cys Leu Gln Leu Ile Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys
145                 150                 155                 160

Gly Ser Tyr Thr Phe Val Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser
                165                 170                 175

Ala Leu Glu Glu Lys Glu Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr
            180                 185                 190

Phe Phe Ile Tyr Gly Gln Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met
        195                 200                 205

Gly His Leu Val Gln Arg Lys Lys Val His Val Phe Gly Asp Glu Leu
    210                 215                 220
```

Ser Leu Val Thr Leu Phe Arg Cys Ile Gln Asn Met Pro Glu Thr Leu
225 230 235 240

Pro Asn Asn Ser Cys Tyr Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly
245 250 255

Asp Glu Leu Gln Leu Ala Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu
260 265 270

Asp Gly Asp Val Thr Phe Phe Gly Ala Leu Lys Leu Leu
275 280 285

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 166..1020

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "nucleotides 3 and 11 are designated C, may be A, C, G,or T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CACAAAGGCA CAAAGGAGAA AATTCAGGAT AACTCTCCTG AGGGGTGAGC CAAGCCCTGC 60

CATGTATTGC ACGCAGGACA TCAACAAACA CAGATAACAG GAAATGATCC ATTCCCTGTG 120

GTCACTTATT CTAAAGGCCC CAACCTTCAA AGTTCAAGTA GTGAT ATG GAT GAC 174
Met Asp Asp
1

TCC ACA GAA AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA 222
Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg
5 10 15

GAA GAA ATG AAA CTG AAG GAG TGT GTT TCC ATC CTC CCA CGG AAG GAA 270
Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro Arg Lys Glu
20 25 30 35

```
AGC CCC TCT GTC CGA TCC TCC AAA GAC GGA AAG CTG CTG GCT GCA ACC      318
Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu Ala Ala Thr
                 40                  45                  50

TTG CTG CTG GCA CTG CTG TCT TGC TGC CTC ACG GTG GTG TCT TTC TAC      366
Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val Ser Phe Tyr
             55                  60                  65

CAG GTG GCC GCC CTG CAA GGG GAC CTG GCC AGC CTC CGG GCA GAG CTG      414
Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg Ala Glu Leu
         70                  75                  80

CAG GGC CAC CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG      462
Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys
     85                  90                  95

GCC GGC CTG GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT      510
Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe
100                 105                 110                 115

GAA CCA CCA GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT      558
Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn
                120                 125                 130

AAG CGT GCC GTT CAG GGT CCA GAA GAA ACA GTC ACT CAA GAC TGC TTG      606
Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Val Thr Gln Asp Cys Leu
            135                 140                 145

CAA CTG ATT GCA GAC AGT GAA ACA CCA ACT ATA CAA AAA GGA TCT TAC      654
Gln Leu Ile Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys Gly Ser Tyr
        150                 155                 160

ACA TTT GTT CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA      702
Thr Phe Val Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu
    165                 170                 175

GAA AAA GAG AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA      750
Glu Lys Glu Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile
180                 185                 190                 195

TAT GGT CAG GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA      798
Tyr Gly Gln Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu
                200                 205                 210

ATT CAG AGG AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG      846
Ile Gln Arg Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val
            215                 220                 225

ACT TTG TTT CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT      894
Thr Leu Phe Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn
        230                 235                 240
```

60

TCC TGC TAT TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC 942
Ser Cys Tyr Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu
245 250 255

CAA CTT GCA ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT 990
Gln Leu Ala Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp
260 265 270 275

GTC ACA TTT TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC 1040
Val Thr Phe Phe Gly Ala Leu Lys Leu Leu
280 285

TGTAGCTATT TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC 1100

C 1101

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asp Asp Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu
1 5 10 15

Lys Lys Arg Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro
20 25 30

Arg Lys Glu Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu
35 40 45

Ala Ala Thr Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val
50 55 60

Ser Phe Tyr Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg
65 70 75 80

Ala Glu Leu Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly
85 90 95

Ala Pro Lys Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu
100 105 110

Lys Ile Phe Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn
115 120 125

Ser Arg Asn Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Val Thr Gln
130 135 140

Asp Cys Leu Gln Leu Ile Ala Asp Ser Glu Thr Pro Thr Ile Gln Lys
145 150 155 160

Gly Ser Tyr Thr Phe Val Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser
165 170 175

Ala Leu Glu Glu Lys Glu Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr
180 185 190

Phe Phe Ile Tyr Gly Gln Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met
195 200 205

Gly His Leu Ile Gln Arg Lys Lys Val His Val Phe Gly Asp Glu Leu
210 215 220

Ser Leu Val Thr Leu Phe Arg Cys Ile Gln Asn Met Pro Glu Thr Leu
225 230 235 240

Pro Asn Asn Ser Cys Tyr Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly
245 250 255

Asp Glu Leu Gln Leu Ala Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu
260 265 270

Asp Gly Asp Val Thr Phe Phe Gly Ala Leu Lys Leu Leu
275 280 285

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 987 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 109..906

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "nucleotides 3, 14, 33, 36, and 43 are designated C, may be A, C, G, or T"

62

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CCCCTTTGGG TCGCCCACTA ATCCGAACAA ACCCCCATAA AACGAAATGA TCCATTCCCT      60

GTGGTCACTT ATTCTAAAGG CCCCAACCTT CAAAGTTCAA GTAGTGAT ATG GAT GAC      117
                                                      Met Asp Asp
                                                       1

TCC ACA GAA AGG GAG CAG TCA CGC CTT ACT TCT TGC CTT AAG AAA AGA      165
Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu Lys Lys Arg
      5                  10                  15

GAA GAA ATG AAA CTG AAG GAG TGT GTT TCC ATC CTC CCA CGG AAG GAA      213
Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro Arg Lys Glu
 20                  25                  30                  35

AGC CCC TCT GTC CGA TCC TCC AAA GAC GGA AAG CTG CTG GCT GCA ACC      261
Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu Ala Ala Thr
                 40                  45                  50

TTG CTG CTG GCA CTG CTG TCT TGC TGC CTC ACG GTG GTG TCT TTC TAC      309
Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val Ser Phe Tyr
             55                  60                  65

CAG GTG GCC GCC CTG CAA GGG GAC CTG GCC AGC CTC CGG GCA GAG CTG      357
Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg Ala Glu Leu
         70                  75                  80

CAG GGC CAC CAC GCG GAG AAG CTG CCA GCA GGA GCA GGA GCC CCC AAG      405
Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly Ala Pro Lys
     85                  90                  95

GCC GGC CTG GAG GAA GCT CCA GCT GTC ACC GCG GGA CTG AAA ATC TTT      453
Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu Lys Ile Phe
100                 105                 110                 115

GAA CCA CCA GCT CCA GGA GAA GGC AAC TCC AGT CAG AAC AGC AGA AAT      501
Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn Ser Arg Asn
                120                 125                 130

AAG CGT GCC GTT CAG GGT CCA GAA GAA ACA GGA TCT TAC ACA TTT GTT      549
Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Gly Ser Tyr Thr Phe Val
            135                 140                 145

CCA TGG CTT CTC AGC TTT AAA AGG GGA AGT GCC CTA GAA GAA AAA GAG      597
Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu Glu Lys Glu
        150                 155                 160

AAT AAA ATA TTG GTC AAA GAA ACT GGT TAC TTT TTT ATA TAT GGT CAG      645
Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile Tyr Gly Gln
    165                 170                 175
```

```
GTT TTA TAT ACT GAT AAG ACC TAC GCC ATG GGA CAT CTA ATT CAG AGG        693
Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu Ile Gln Arg
180                 185                 190                 195

AAG AAG GTC CAT GTC TTT GGG GAT GAA TTG AGT CTG GTG ACT TTG TTT        741
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
                200                 205                 210

CGA TGT ATT CAA AAT ATG CCT GAA ACA CTA CCC AAT AAT TCC TGC TAT        789
Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn Ser Cys Tyr
            215                 220                 225

TCA GCT GGC ATT GCA AAA CTG GAA GAA GGA GAT GAA CTC CAA CTT GCA        837
Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu Gln Leu Ala
        230                 235                 240

ATA CCA AGA GAA AAT GCA CAA ATA TCA CTG GAT GGA GAT GTC ACA TTT        885
Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp Val Thr Phe
    245                 250                 255

TTT GGT GCA TTG AAA CTG CTG TGACCTACTT ACACCATGTC TGTAGCTATT           936
Phe Gly Ala Leu Lys Leu Leu
260                 265

TTCCTCCCTT TCTCTGTACC TCTAAGAAGA AAGAATCTAA CTGAAAATAC C               987
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 266 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asp Asp Ser Thr Glu Arg Glu Gln Ser Arg Leu Thr Ser Cys Leu
 1               5                  10                  15

Lys Lys Arg Glu Glu Met Lys Leu Lys Glu Cys Val Ser Ile Leu Pro
            20                  25                  30

Arg Lys Glu Ser Pro Ser Val Arg Ser Ser Lys Asp Gly Lys Leu Leu
        35                  40                  45

Ala Ala Thr Leu Leu Leu Ala Leu Leu Ser Cys Cys Leu Thr Val Val
    50                  55                  60

Ser Phe Tyr Gln Val Ala Ala Leu Gln Gly Asp Leu Ala Ser Leu Arg
65                  70                  75                  80
```

Ala Glu Leu Gln Gly His His Ala Glu Lys Leu Pro Ala Gly Ala Gly
85 90 95

Ala Pro Lys Ala Gly Leu Glu Glu Ala Pro Ala Val Thr Ala Gly Leu
100 105 110

Lys Ile Phe Glu Pro Pro Ala Pro Gly Glu Gly Asn Ser Ser Gln Asn
115 120 125

Ser Arg Asn Lys Arg Ala Val Gln Gly Pro Glu Glu Thr Gly Ser Tyr
130 135 140

Thr Phe Val Pro Trp Leu Leu Ser Phe Lys Arg Gly Ser Ala Leu Glu
145 150 155 160

Glu Lys Glu Asn Lys Ile Leu Val Lys Glu Thr Gly Tyr Phe Phe Ile
165 170 175

Tyr Gly Gln Val Leu Tyr Thr Asp Lys Thr Tyr Ala Met Gly His Leu
180 185 190

Ile Gln Arg Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val
195 200 205

Thr Leu Phe Arg Cys Ile Gln Asn Met Pro Glu Thr Leu Pro Asn Asn
210 215 220

Ser Cys Tyr Ser Ala Gly Ile Ala Lys Leu Glu Glu Gly Asp Glu Leu
225 230 235 240

Gln Leu Ala Ile Pro Arg Glu Asn Ala Gln Ile Ser Leu Asp Gly Asp
245 250 255

Val Thr Phe Phe Gly Ala Leu Lys Leu Leu
260 265

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 617 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..452

65

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
TC TCA GCT CCT CCT GCA CCA TGC CTG CCT GGA TGC CGC CAT TCT CAA          47
   Ser Ala Pro Pro Ala Pro Cys Leu Pro Gly Cys Arg His Ser Gln
     1               5                  10                  15

CAT GAT GAT AAT GGA ATG AAC CTC AGA AAC AGA ACT TAC ACA TTT GTT          95
His Asp Asp Asn Gly Met Asn Leu Arg Asn Arg Thr Tyr Thr Phe Val
                 20                  25                  30

CCA TGG CTT CTC AGC TTT AAA AGA GGA AAT GCC TTG GAG GAG AAA GAG         143
Pro Trp Leu Leu Ser Phe Lys Arg Gly Asn Ala Leu Glu Glu Lys Glu
             35                  40                  45

AAC AAA ATA GTG GTG AGG CAA ACA GGC TAT TTC TTC ATC TAC AGC CAG         191
Asn Lys Ile Val Val Arg Gln Thr Gly Tyr Phe Phe Ile Tyr Ser Gln
         50                  55                  60

GTT CTA TAC ACG GAC CCC ATC TTT GCT ATG GGT CAT GTC ATC CAG AGG         239
Val Leu Tyr Thr Asp Pro Ile Phe Ala Met Gly His Val Ile Gln Arg
     65                  70                  75

AAG AAA GTA CAC GTC TTT GGG GAC GAG CTG AGC CTG GTG ACC CTG TTC         287
Lys Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe
 80                  85                  90                  95

CGA TGT ATT CAG AAT ATG CCC AAA ACA CTG CCC AAC AAT TCC TGC TAC         335
Arg Cys Ile Gln Asn Met Pro Lys Thr Leu Pro Asn Asn Ser Cys Tyr
                100                 105                 110

TCG GCT GGC ATC GCG AGG CTG GAA GAA GGA GAT GAG ATT CAG CTT GCA         383
Ser Ala Gly Ile Ala Arg Leu Glu Glu Gly Asp Glu Ile Gln Leu Ala
            115                 120                 125

ATT CCT CGG GAG AAT GCA CAG ATT TCA CGC AAC GGA GAC GAC ACC TTC         431
Ile Pro Arg Glu Asn Ala Gln Ile Ser Arg Asn Gly Asp Asp Thr Phe
        130                 135                 140

TTT GGT GCC CTA AAA CTG CTG TAACTCACTT GCTGGAGTGC GTGATCCCCT            482
Phe Gly Ala Leu Lys Leu Leu
    145                 150

TCCCTCGTCT TCTCTGTACC TCCGAGGGAG AAACAGACGA CTGGAAAAAC TAAAAGATGG       542

GGAAAGCCGT CAGCGAAAGT TTTCTCGTGA CCCGTTGAAT CTGATCCAAA CCAGGAAATA       602

TAACAGACAG CCACA                                                       617
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Ser Ala Pro Pro Ala Pro Cys Leu Pro Gly Cys Arg His Ser Gln His
 1               5                  10                  15

Asp Asp Asn Gly Met Asn Leu Arg Asn Arg Thr Tyr Thr Phe Val Pro
            20                  25                  30

Trp Leu Leu Ser Phe Lys Arg Gly Asn Ala Leu Glu Glu Lys Glu Asn
        35                  40                  45

Lys Ile Val Val Arg Gln Thr Gly Tyr Phe Phe Ile Tyr Ser Gln Val
    50                  55                  60

Leu Tyr Thr Asp Pro Ile Phe Ala Met Gly His Val Ile Gln Arg Lys
65                  70                  75                  80

Lys Val His Val Phe Gly Asp Glu Leu Ser Leu Val Thr Leu Phe Arg
                85                  90                  95

Cys Ile Gln Asn Met Pro Lys Thr Leu Pro Asn Asn Ser Cys Tyr Ser
            100                 105                 110

Ala Gly Ile Ala Arg Leu Glu Glu Gly Asp Glu Ile Gln Leu Ala Ile
        115                 120                 125

Pro Arg Glu Asn Ala Gln Ile Ser Arg Asn Gly Asp Asp Thr Phe Phe
    130                 135                 140

Gly Ala Leu Lys Leu Leu
145                 150
```

WHAT IS CLAIMED IS:

1. An isolated or recombinant polypeptide which:

(a) comprises a plurality of epitopes found on; and

(b) exhibits at least 85% sequence identity over a length of at least 12 contiguous amino acids to;

a polypeptide selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 and SEQ ID NO: 8.

2. A polypeptide of Claim 1, wherein the polypeptide binds with specificity to an antibody generated against an immunogen selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 and SEQ ID NO: 8.

3. A fusion protein comprising a polypeptide according to either Claim 1 or 2.

4. A composition comprising a polypeptide or fusion protein according to any one of Claims 1 to 3.

5. An antibody which specifically binds a polypeptide of either Claim 1 or 2.

6. An isolated nucleic acid encoding a polypeptide or fusion protein according to any one of Claims 1 to 3.

7. A recombinant vector comprising the nucleic acid of Claim 6.

8. A host cell comprising the recombinant vector of Claim 7.

9. A method for producing a polypeptide or fusion protein comprising culturing the host cell of Claim 8 under conditions in which the nucleic acid is expressed.

69

10. A kit comprising:

a) a polypeptide or fusion protein according to any one of Claims 1 to 3;

b) an antibody which specifically binds to a polypeptide according to either Claim 1 or 2; or

c) a nucleic acid according to Claim 6.